Conformational changes involved in signal transduction in avian cryptochromes



Angela Sonia Gehrckens

Corpus Christi College, Oxford

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Physical and Theoretical Chemistry

Trinity Term 2023

Acknowledgements

There are so many people who I would like to thank and acknowledge who have walked alongside me for a bit during these past four years and without whom it would not have been the same experience.

First and foremost, I would like to thank my supervisors Prof. Peter J. Hore and Prof. Justin L. P. Benesch who offered me the opportunity to do my DPhil with them and without whom I would not have had such an interesting topic to work on. Every discussion with both Peter and Justin was always very helpful and led to new ideas or a better understanding of the problem at hand.

I would also like to thank Prof. Christiane R. Timmel, who half-adopted me into her group to learn about EPR, and Prof. Stuart R. Mackenzie, who has equally supported me as if he were another supervisor.

Thanks to Patrick who was always there, from the moment we were awkwardly waiting to receive our new computers, to the 100th ATS sandwich, to the 30th movie discussion that featured a recommendation of Wall-E, and to the last helpful text message. Thanks to Shane for introducing me into the world of HDX-MS and thanks to Phil for his support during my first year. I would also like to thank Ilia Solov'yov who taught me about MD simulations as well as Fabian Schuhmann and Maja Hanić who helped me with all MD related questions and who I got to spend some time with over the years. I am very happy that I was able to meet all of the 'bird people'; thanks to Dan for always making sure I had lunch and working with me on my pronunciation (Schroppschirre!), Jamie for the understanding, the caring questions and the wine, Hang for the interesting conversations, Tommy for the chocolate and the last-minute EPR help when I needed it most, Ana for more chocolate, her understanding and caring, Kevin for storing, sorting and forwarding an unknown but massive amount of protein samples for me and for always being happy to help whenever he could, Ashley for helpful recommendations and extremely good food, Pip for always believing that the HDX might be working, Carl for unbreakable happiness that he shares with everyone, Laurynas for reminding me why topics that started to frustrate me might actually be interesting, and to Lewis for the ice cream to counteract the saaadness and the music recommendations. Thanks to Olivier for teaching me a really hard and valuable lesson, but also for listening.

Massive thanks go out to Laila who was the first person I could discuss absolutely every little detail about my work with and who has spent endless hours working

together with me on the same project. My PhD would not have been the same without her and her helpful ideas, her enthusiasm and her drive to bring the project forward. Thanks also for all the tea and the time we spent at conferences together!

I am also thankful to have met everyone in Justin's and Carol's groups and I would like to thank Dom for his support and the long walks during the pandemic, Wiktoria for her enthusiasm and a fun pandemic weekend in Germany, Luke, Josh, Guodong, Ellie, Fran, Zihao, Valeria, Tarick (thanks as well for all the help with the crosslinking!), Corinne, John, and Abraham for many helpful conversations over the years, Neel for all the MD discussions and interesting conversations, Joana for the walks and conversations, Olav for lots of tea, strange walls, and good conversations, and Konstantin for sharing useful EPR knowledge. Thanks also to Virginia for listening and helping when most needed.

I would also like to thank the people who have taught me about EPR at the instruments: Will, Tommy (again!), and especially Arnau who spent hours explaining and showing things to me.

Thanks a lot to Sabine, Andy, Vicky, Parisa, Rachel, Carolyn, and Paul who made all the organisational parts of my DPhil (including fixing or replacing lots of broken things) so much easier.

I would also like to thank a number of people who I have not directly been working with but who I had the pleasure of spending some valuable time with. Thanks to Phil for the conversations and for never forgetting my birthday, Thomas for interesting conversations about life, all the people at Corpus Christi who made my time at this College so much more enjoyable, Rita for being incredibly supportive and always there when I needed a friend, Chloe and Guillem for French food and fun evenings at the pub, Luca for having been such a motivator since I read that book about him by pure chance, Ekin for insightful conversations, Andy for listening, Jonathan for many cold winter walks, all my housemates for almost driving me insane, Tjark for being a friend and the trip to the festival, Ron for picking me up whenever necessary and for being a great human as well as a true friend, Sofie for all the laughs, the motivation and countless adventures since back in undergrad, and Tom for his support, for listening and for all the fun moments despite all the odds. I am incredibly happy I met you here.

Thanks to my family for always believing in me.

Abstract

It is hypothesised that the putative magnetoreceptor protein cryptochrome enables night-migratory birds to navigate using the Earth's magnetic field in a light-dependent way. Following a well-established radical pair mechanism, different downstream chemical and biochemical reactions are thought to take place depending on how the inclination of the Earth's magnetic field influences the singlet and triplet yields of the magnetically sensitive radical pair. This is thought to lead to protein conformational changes followed by protein-protein interactions with interaction partners.

These conformational changes and interactions have been investigated as the first steps of the magnetoreception signalling process using native mass spectrometry, mass photometry, chemical crosslinking followed by mass spectrometric analysis, and hydrogen-deuterium exchange mass spectrometry (HDX-MS). The latter technique is capable of discovering interaction surfaces and conformational changes by comparing surface exposures. Recently, it was possible to detect conformational changes in certain protein regions of cryptochrome 4a of several bird species using this method, with especially prominent changes found in the phosphate-binding loop, the protrusion motif, and the C-terminal tail. Differences were seen when the experiments were performed with and without blue light. The results are detailed in chapters 3 and 4. Small changes due to an applied static magnetic field were investigated in detail in chapter 4.

Since it was recently found that avian cryptochromes can form dimers and as dimers of non-bird cryptochromes seem to play important roles in nature, chemical crosslinking was performed to narrow down the interaction surface between monomers in potentially functional dimeric cryptochrome structures, complemented by molecular dynamics simulations performed by collaborating research groups.

Two residues, C317 and C412, were found to play a role in light-induced dimerisation via disulphide bonding. To further investigate the interaction surface, electron paramagnetic resonance spectroscopy was used, more specifically double electron-electron resonance, to complement the previously acquired data with cysteine-residue-dependent distance measurements using nitroxide spin labels. The findings on dimerisation are presented in chapter 5.

Finally, by mapping conformational changes found in molecular dynamics simulations onto changes observed in a previous HDX-MS data set on the *Drosophila*

cryptochrome, a few gaps left by the incomplete coverage of the current experimental data were filled by acquiring information on the phosphate-binding loop and the protrusion motif, two regions which have also been found to undergo light-induced structural changes for the avian cryptochromes investigated. The results of this theoretical investigation are discussed in chapter 6.

Using a range of complementary techniques, this thesis aims to elucidate protein conformational changes as the first step in signal transduction in several cryptochrome species finding similar changes in the same loop regions across multiple proteins, tackled the question whether a functional dimer might exist in the context of magnetoreception and found two important residues involved in dimerisation, and lastly used the finding that this dimerisation process is light-induced to start investigating the interaction surfaces in these dimeric structures.

Statement of Originality

This thesis describes research conducted in the Department of Chemistry at the University of Oxford between October 2019 and June 2023. The individual chapters focus on the research theme of elucidating conformational changes and dynamics of different cryptochrome species, a protein thought to be involved in magnetoreception in birds. All data presented were acquired and analysed by me unless otherwise indicated in the main text or in the appropriate figure captions. All figures were likewise generated by me, unless otherwise indicated. The results presented in chapter 5 resulted from a highly collaborative effort across multiple research groups and the contributions have been indicated accordingly. This work has not been submitted for a degree, diploma, or any other academic certification at any other institution, in whole or in part.

Angela S. Gehrckens
October 2023

Contents

List of Figures **xiii**

List of Abbreviations **xvii**

1 Introduction **1**
1.1 Motivation . . . 1
1.2 The geomagnetic field . . . 2
1.3 Magnetoreception . . . 4
1.4 Avian navigation . . . 5
1.4.1 Different theories of avian navigation . . . 7
1.5 The radical pair mechanism . . . 9
1.6 Cryptochromes . . . 15
1.6.1 Cryptochrome's involvement in the RPM . . . 19
1.6.2 Cryptochrome's C-terminal tail . . . 20
1.6.3 Cryptochrome oligomerisation . . . 23
1.7 Structural investigation of light-sensitive proteins using HDX-MS . 24
1.8 Contribution . . . 26
References . . . 27

2 Technical Background **43**
2.1 Mass spectrometry and mass photometry . . . 43
2.1.1 Mass spectrometry instrumentation . . . 44
2.1.2 Solution hydrogen/deuterium exchange mass spectrometry . 55
2.1.3 Mass photometry . . . 62
2.2 Electron Paramagnetic Resonance Spectroscopy . . . 64
2.2.1 Continuous wave EPR . . . 65
2.2.2 Pulse EPR . . . 66
2.3 Molecular Dynamics Simulations . . . 69
2.3.1 Force fields and water models . . . 71
2.3.2 Constraints and conditions: periodic boundary conditions, cutoffs, thermo- and barostats . . . 73
References . . . 74

3 Conformational changes through the lens of HDX-MS 87
3.1 Introduction . . . 87
3.2 Light-induced conformational changes in European robin cryptochrome 89
3.3 Introducing the *Er*Cry4a W369F mutant . . . 93
3.3.1 Could different buffer conditions improve our results? . . . 101
3.3.2 Tracking down the C-terminal tail . . . 104
3.3.3 Concluding discussion . . . 107
3.3.4 Final remarks . . . 109
3.4 Material and Methods . . . 110
3.4.1 Material . . . 110
3.4.2 HDX-MS . . . 111
3.4.3 Native MS . . . 113
3.4.4 Data analysis . . . 113
References . . . 114

4 A comparative study of conformational changes within different avian cryptochrome species 119
4.1 Introduction . . . 120
4.2 Light-induced conformational changes of avian cryptochromes . . . 122
4.2.1 Three different avian species . . . 122
4.2.2 Uptake comparison . . . 125
4.2.3 Statistical data analysis . . . 130
4.2.4 Conclusions . . . 164
4.3 Outlook and Interaction partner studies . . . 167
4.4 Materials and Methods . . . 169
4.4.1 Materials . . . 169
4.4.2 HDX-MS . . . 169
4.4.3 Native MS . . . 170
4.4.4 Data analysis . . . 171
References . . . 171

5 Dimerisation of *Er*Cry4a 175
5.1 Introduction . . . 176
5.2 Results and Discussion . . . 176
5.2.1 Covalent dimerisation of *Er*Cry4a . . . 176
5.2.2 Chemical cross-linking followed by mass spectrometric analysis 183
5.2.3 Focus on C317 and C412 . . . 184
5.2.4 Dimerisation under different light conditions . . . 186
5.2.5 The search for the interaction interface . . . 192
5.3 Conclusion . . . 201

5.4 Materials and Methods . . . 203
5.4.1 Materials . . . 203
5.4.2 Native MS . . . 203
5.4.3 Mass photometry . . . 204
5.4.4 SDS-PAGE . . . 204
5.4.5 Chemical cross-linking and mass spectrometric analysis . . . 205
5.4.6 DEER . . . 206
References . . . 207

6 Complementing experimental structural data through molecular dynamics simulations 213
6.1 Introduction . . . 213
6.2 Supporting HDX-MS data on *Dm*Cry . . . 215
6.2.1 Results and Discussion . . . 215
6.2.2 Conclusion . . . 220
6.3 Methods . . . 221
References . . . 222

7 Conclusions and future work 227
References . . . 230

Appendices

A Appendix 235
A.1 Appendix for chapter 3 . . . 236
A.1.1 Peptide coverage for *Er*Cry4a WT and *Er*Cry4a W369F. . . . 236
A.1.2 Uptake curves for *Er*Cry4a W369F peptides of interest. . . . 237
A.1.3 Uptake curves on a logarithmic time scale for *Er*Cry4a W369F peptides of interest. . . . 241
A.2 Appendix for chapter 4 . . . 245
A.2.1 Relative fractional uptake . . . 245
A.2.2 Chiclet plots created in HD-eXplosion . . . 249
A.2.3 ΔRFU plots created by Laila M. N. Shah . . . 256
A.3 Appendix for chapter 5 . . . 271

List of Figures

1.1 The Earth's magnetic field. . . . 3
1.2 Simple reaction scheme of the radical pair mechanism. . . . 11
1.3 Proposed photochemical scheme for photo-activation and and radical pair formation in cryptochromes. . . . 13

2.1 Electrospray ionisation of proteins. . . . 48
2.2 Quadrupole mass analyser. . . . 51
2.3 Quadrupole time-of-flight and orbitrap mass analysers. . . . 53
2.4 The common 'bottom-up' HDX-MS workflow. . . . 57
2.5 HDX-MS kinetics: EX1 &EX2. . . . 62
2.6 Working principle of mass photometry. . . . 63
2.7 Pulse sequence for 4-pulse DEER. . . . 68

3.1 Relative fractional uptake of *Er*Cry4a WT. . . . 91
3.2 Relative uptake of representative *Er*Cry4a WT peptides. . . . 92
3.3 Relative fractional uptake of *Er*Cry4a W369F. . . . 94
3.4 Relative fractional uptake of *Er*Cry4a W369F displayed on AlphaFold structures. . . . 95
3.5 Relative uptake of representative *Er*Cry4a W369F peptides. . . . 96
3.6 Native mass spectra of *Er*Cry4a WT and *Er*Cry4a W369F. . . . 102
3.7 Relative uptake of representative *Cc*Cry4a WT peptides at different buffer conditions. . . . 103
3.8 Results of a D_2O concentration trial. . . . 106
3.9 Magnet strength dependence on distance. . . . 111
3.10 Emission spectra of all lights used during experiments. . . . 113

4.1 Relative fractional uptake of *Er*Cry4a WT and coverage maps. . . . 123
4.2 Native mass spectra of seven avian cryptochrome 4a species. . . . 124
4.3 SDS-PAGE gel of *Gg*Cry4a W369F and *Cc*Cry4a W369F at different HDX-MS experimental conditions. . . . 125
4.4 Relative fractional uptake of *Er*Cry4a WT. . . . 127
4.5 Relative fractional uptake of *Er*Cry4a W369F. . . . 128
4.6 Relative fractional uptake of *Er*Cry4a R356E. . . . 129

4.7 Deuteros Woods plot of *Er*Cry4a WT, LS-DS. 131
4.8 Peptides of significant uptake of *Er*Cry4a WT, W369F, and R356E for the LS-DS comparison. 144
4.9 Example significant peptides found in hdxstats. 157

5.1 Dimerisation in *Er*Cry4a. 177
5.2 Close up on dimerisation. 179
5.3 Solvent accessible cysteine residues. 180
5.4 Network of cross-links. 184
5.5 Focus on C317 and C412. 186
5.6 Denaturing SDS-PAGE gels monitoring dimerisation over time in the dark and under blue-light conditions of different *Er*Cry4a mutants. 187
5.7 Further SDS-PAGE gels monitoring dimerisation over time in the dark and under blue-light conditions of different *Er*Cry4a mutants on denaturing SDS-PAGE gels. 190
5.8 Extent of dimerisation of different avian cryptochrome species on denaturing SDS-PAGE gels. 191
5.9 Cysteines of *Er*Cry4a WT. 193
5.10 Native MS of IPSL-labelled *Er*Cry4a 5-Cys. 195
5.11 Electron spin-echo field-swept spectra. 196
5.12 Distances obtained for *Er*Cry4a 5-Cys using DEER (first experiment). 197
5.13 Distances obtained for unlabelled *Er*Cry4a 5-Cys using DEER. . . . 198
5.14 Distances obtained for *Er*Cry4a 5-Cys using DEER (second experiment). 199

6.1 Conformational changes observed in molecular dynamics simulations. 216
6.2 Conformational dynamics of *Dm*Cry WT observed by HDX-MS. . . 218

A.1 Coverage maps for *Er*Cry4a WT and *Er*Cry4a W369F. 236
A.2 Relative uptake of *Er*Cry4a W369F peptides of interest. 237
A.3 Relative uptake of *Er*Cry4a W369F peptides of interest. 238
A.4 Relative uptake of *Er*Cry4a W369F peptides of interest. 239
A.5 Relative uptake of *Er*Cry4a W369F peptides of interest. 240
A.6 Relative uptake of *Er*Cry4a W369F peptides of interest. 241
A.7 Relative uptake of *Er*Cry4a W369F peptides of interest. 242
A.8 Relative uptake of *Er*Cry4a W369F peptides of interest. 243
A.9 Relative uptake of *Er*Cry4a W369F peptides of interest. 244
A.10 Relative fractional uptake of *Cc*Cry4a WT. 245
A.11 Relative fractional uptake of *Cc*Cry4a W369F. 246
A.12 Relative fractional uptake of *Gg*Cry4a WT. 247

A.13 Relative fractional uptake of *Gg*Cry4a W369F. 248
A.14 HD-eXplosion chiclet plots of *Er*Cry4a WT significant peptides. . . 249
A.15 HD-eXplosion chiclet plots of *Er*Cry4a W369F significant peptides. 250
A.16 HD-eXplosion chiclet plots of *Er*Cry4a R356E significant peptides. . 251
A.17 HD-eXplosion chiclet plots of *Cc*Cry4a WT significant peptides. . . 252
A.18 HD-eXplosion chiclet plots of *Cc*Cry4a W369F significant peptides. 253
A.19 HD-eXplosion chiclet plots of *Gg*Cry4a WT significant peptides. . . 254
A.20 HD-eXplosion chiclet plots of *Gg*Cry4a W369F significant peptides. 255
A.21 Relative fractional uptake of *Er*Cry4a WT and W369F significant peptides. 256
A.22 Relative fractional uptake of *Er*Cry4a WT and W369F significant peptides after 15 s of labelling time. 257
A.23 Relative fractional uptake of *Er*Cry4a WT and W369F significant peptides after 1 min of labelling time. 258
A.24 Relative fractional uptake of *Er*Cry4a WT and W369F significant peptides after 5 min of labelling time. 259
A.25 Relative fractional uptake of *Er*Cry4a WT and W369F significant peptides after 30 min of labelling time. 260
A.26 Relative fractional uptake of *Cc*Cry4a WT and W369F significant peptides. 261
A.27 Relative fractional uptake of *Cc*Cry4a WT and W369F significant peptides after 15 s of labelling time. 262
A.28 Relative fractional uptake of *Cc*Cry4a WT and W369F significant peptides after 1 min of labelling time. 263
A.29 Relative fractional uptake of *Cc*Cry4a WT and W369F significant peptides after 5 min of labelling time. 264
A.30 Relative fractional uptake of *Cc*Cry4a WT and W369F significant peptides after 30 min of labelling time. 265
A.31 Relative fractional uptake of *Gg*Cry4a WT and W369F significant peptides. 266
A.32 Relative fractional uptake of *Gg*Cry4a WT and W369F significant peptides after 15 s of labelling time. 267
A.33 Relative fractional uptake of *Gg*Cry4a WT and W369F significant peptides after 1 min of labelling time. 268
A.34 Relative fractional uptake of *Gg*Cry4a WT and W369F significant peptides after 5 min of labelling time. 269
A.35 Relative fractional uptake of *Gg*Cry4a WT and W369F significant peptides after 30 min of labelling time. 270
A.36 Structures of the non-covalently bound dimers ncov^A and ncov^M. . 271

A.37 Structures of the covalently bound dimers of *Er*Cry4a, cov^{317A} and cov^{317B}. 271
A.38 Structure of the doubly covalently bound *Er*Cry4a cov^{D} dimer. . . 272
A.39 Structures of the two favourable *Er*Cry4a dimers cov^{412A} and cov^{412B}. 273
A.40 Echo intensity of a $FADH^{\bullet}$ radical during magnetic field sweep. . . . 274
A.41 Electron spin-echo field-swept spectra. 274

List of Abbreviations

ACN . . . . . Acetonitrile

At . . . . . *Arabidopsis thaliana*

ATP . . . . . Adenosine triphosphate

BLUF . . . . . Blue light using FAD

BME . . . . . *beta*-mercaptoethanol

Cc . . . . . *Coturnix coturnix*

CI . . . . . Confidence interval

CID . . . . . Collision-induced dissociation

Cl . . . . . *Columba livia*

Cra . . . . . *Chlamydomonas reinhardtii*

CRP . . . . . C-reactive protein

Cry . . . . . Cryptochrome

CTT . . . . . C-terminal tail

cw . . . . . Continuous wave

DC . . . . . Direct current

DEER . . . . . Double electron-electron resonance

Dm . . . . . *Drosophila melanogaster*

DS . . . . . Dark state

DSBU . . . . . Disuccinimidyl dibutyric acid

DSM . . . . . Dark state + magnet

DSSO . . . . . Disuccinimidyl sulphoxide

DTNB . . . . . 5,5-dithio-bis-(2-nitrobenzoic acid) (Ellman's reagent)

DTT . . . . . Dithiothreitol

ECD . . . . . Electron-capture dissociation

EPR . . . . . Electron paramagnetic resonance spectroscopy

Er . . . . . *Erithacus rubecula*

ESR . . . . . Electron spin resonance spectroscopy

ETD . . . . . Electron-transfer dissociation

FAD . . . . . Flavin adenine dinucleotide

FDR . . . . . False discovery rate

GdnHCl . . . . Guanidinium hydrochloride

Gg . . . . . *Gallus gallus*

GNAT2 . . . . Guanine nucleotide-binding protein G(t) subunit α-2

HCD . . . . . higher-energy C-trap dissociation / higher energy collisional dissociation

HDX . . . . . hydrogen/deuterium exchange

IPSL . . . . . (Iodoacetamido)proxyl

LC . . . . . Liquid chromatography

LeuEnk . . . . Leucine enkephalin, peptide of amino acid sequence Tyr-Gly-Gly-Phe-Leu

LJ . . . . . Lennard-Jones

LOV . . . . . Light-oxygen-voltage

LS . . . . . Light state

LSM . . . . . Light state + magnet

m/z . . . . . Mass over charge ratio

MC . . . . . Monte Carlo

MD . . . . . Molecular dynamics

MS . . . . . mass spectrometry

MS/MS . . . . Tandem mass spectrometry

MTHF . . . . Methenyltetrahydrofolate

(n)ESI . . . . . (nano-)Electrospray ionisation

NMR . . . . . Nuclear magnetic resonance

PBC . . . . . Periodic boundary conditions

PBL . . . . . Phosphate-binding loop

PCA . . . . . Principal component analysis

PDB . . . . . Protein data bank

PELDOR . . . Pulsed electron double resonance

PHR . . . . . Photolyase homology region

PM . . . . . Protrusion motif

PME . . . . . Particle mesh Ewald summation

QM/MM . . . Quantum mechanics / molecular mechanics

(Q-)ToF . . . . (Quadrupole) Time-of-flight

rf . . . . . radio-frequency

RFU . . . . . Relative fractional uptake

RMSD . . . . . root-mean-square deviation

RMSF . . . . . root-mean-square fluctuation

RPM . . . . . Radical pair mechanism

SASA . . . . . Solvent accessible surface area

SDS-PAGE . . sodium dodecyl sulphate–polyacrylamide gel electrophoresis

SEC . . . . . Size exclusion chromatography

UVPD . . . . . Ultra-violet photo dissociation

WT . . . . . Wild type

XL-MS . . . . Chemical cross-linking followed by mass spectrometric analysis

It's part of nature. It's just a part of nature we don't yet understand.

— Rupert Sheldrake in *How Do Birds Find Their Way Home? | Paranormal Pigeons*

1 Introduction

Contents

1.1 Motivation **1**
1.2 The geomagnetic field **2**
1.3 Magnetoreception **4**
1.4 Avian navigation **5**
1.4.1 Different theories of avian navigation 7
1.5 The radical pair mechanism **9**
1.6 Cryptochromes **15**
1.6.1 Cryptochrome's involvement in the RPM 19
1.6.2 Cryptochrome's C-terminal tail 20
1.6.3 Cryptochrome oligomerisation 23
1.7 Structural investigation of light-sensitive proteins using HDX-MS **24**
1.8 Contribution **26**
References **27**

1.1 Motivation

Many species on Earth use a variety of different tricks to navigate with some of them fulfilling navigational tasks that have left humans in awe for centuries. Some bird species are of special interest as they navigate for thousands of kilometres,

often seemingly without any prior knowledge of their destination or guidance on how to reach it [1]. There are several theories tackling the question of which senses might enable migratory birds to perform these tasks [2–4] with the theory that they perceive and use the Earth's magnetic field as one of the most popular ones [2].

Research has been looking into this subject for decades with one class of proteins having been suggested as the most likely putative magnetoreceptor in migratory birds about two decades ago [5]. Currently, cryptochromes are the most likely candidates to fulfil this purpose and have been the subject of intense study. A wealth of interdisciplinary research has accumulated on this topic over the years, ranging from a theoretical analysis of how the Earth's magnetic field might be perceived at an atomistic level involving calculations of electron transfer pathways to the changes observed at the level of tissues or whole organisms for example by tracking migrating birds [2, 6].

However, little is known to date about how exactly cryptochromes might encode directional information after perception of the Earth's magnetic field and about how the perceived information might be forwarded to the brain. To address these questions, this work will make a start to investigate protein conformational changes of various cryptochromes upon exposure to different environmental conditions, tackle the question whether dimerisation might have a biological function and try to fill in the gaps in the experimental data by simulating protein motion.

1.2 The geomagnetic field

Our planet is essentially one enormous magnet with its poles near the geographic poles and a resulting magnetic field [7]. It is covered in magnetic field lines running from the South to the North with inclinations angles ranging from -90 to +90 degrees with horizontal field lines at 0 degrees at the equator as roughly displayed in Figure 1.1. The field strength also changes with the position above the planet. Despite the presence of natural or human-induced distortions to the magnetic field, it still provides a reliable navigational source for orientation as the observed distortions of the field are minimal in comparison to its regularity [8–10].

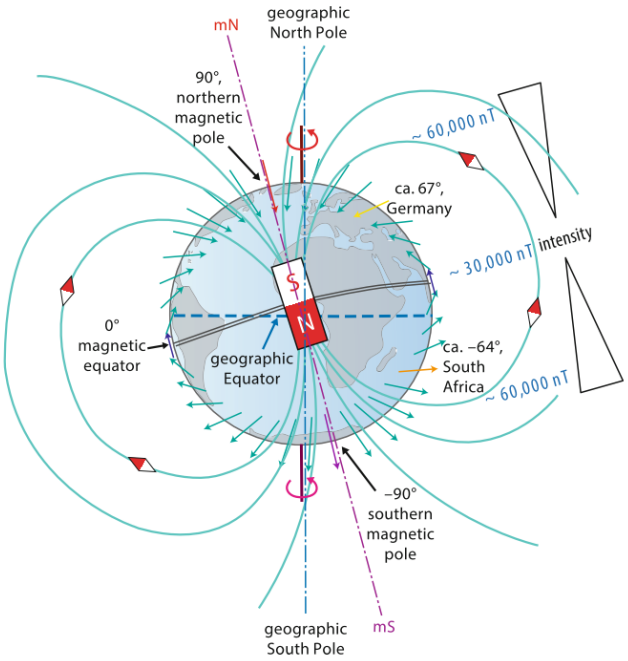

Figure 1.1: The Earth's magnetic field. The Earth's magnetic field lines run from the Southern to the Northern part of the Earth. The intensity of the magnetic field changes from North to South as does the inclination angle. Adapted from Mouritsen (2013) [11].

Moving electronic charges produce a magnetic field. It is usually measured in Tesla, a unit that describes the magnetic flux density [9, 11]. However, the most accepted theory explaining how the Earth's magnet field comes into existence and is maintained is a little more complicated than this and called the geodynamo. Put very simply, this rather complex theory explains how the Earth's magnet field originates from electric currents from within the liquid outer core of the Earth rather than from a permanent magnet [12, 13].

The Earth's magnetic field can change unpredictably over the course of months and years. These variations can for example stem from magnetic storms which are e.g. caused by sun spot activity, or slow drifts in declination and inclination over the course of several years. However, the latter are generally too slow to affect a single organism within its lifespan. Daily changes can also occur, caused by the solar electromagnetic radiation causing currents in the upper atmosphere and ionosphere of the sun-facing side of the planet [14]. The Earth's magnetic field

strength currently ranges from about 24 - 66 µT and is dependent mostly on the latitude [15, 16]. On average, the field changes only by about 3 nT km^{-1} and 0.009° km^{-1} on the north-south axis and even less on the east-west axis [17].

The Earth's magnetic field lines could theoretically both be used as a compass and a map for navigational purposes. Both would be necessary in species that are true navigators and therefore able to find their way after displacement to an unknown area [6].

1.3 Magnetoreception

Many animals have been shown to use the Earth's magnetic field for navigational purposes. Examples are turtles that have an inherent map of where their destination is located without ever having migrated before (probably making use of the Earth's magnetic field lines' inclination and intensity to navigate) [18] or birds, more precisely passerines, that use a magnetic compass amongst other methods with some of them also having a map sense used for homing [19].

It has been shown that a multitude of effects in biological systems can be caused by the addition, change or reduction of surrounding weak magnetic fields [15]. For example, Giorgi *et al.* found increased body size of *Drosophila melanogaster* under constant exposure to magnetic fields of 0.4-0.7 mT [20]. Another interesting study conducted by Ahmad *et al.* showed that plant hypocotyl growth is influenced by the strength of the surrounding magnetic field under blue-light conditions [21]. However, unfortunately these findings could not be verified in an independent study [22].

Aside from pure exposure experiments, conditioning is a tool often used for behavioural experiments in the field of sensory physiology. However, they seem to be riddled with difficulties when it comes to magnetoreception and are rarely successfully repeated [23–25]. However, to date a variety of organisms have been reported to respond to an external magnetic field, such as honey bees [26], sea turtles [27], fish [28], monarch butterflies [29], bats [30], and amphibians (more precisely salamanders) [31], of which the latter have been the first example of a species shown to have a light-dependent magnetic compass sense. Even humans have been

suggested to have the capability to perceive the Earth's magnetic field despite being believed to not have a magnetic sense [32, 33]. However, a class of animals that has been studied extensively in behavioural experiments are birds [34–38] which is probably no wonder given their astonishing abilities to both fly and navigate over long distances since long before humans managed to achieve the same feat.

1.4 Avian navigation

A lot of research has been performed on avian navigation, investigating different methods of navigation and different bird species, often finding that the exact mode of navigation is species-dependent. A sizeable amount of research has been done on night-migratory songbirds with an important result being that these animals seem to have three phases of navigation, the long-distance phase, the homing phase and the pinpointing-the-goal phase. During the long-distance phase the birds make use of celestial and magnetic cues and the last two phases depend on multisensory maps and landmarks [17]. As the work presented in this thesis focuses on how long-distance avian navigation might be performed, this sub-chapter will have the same focus. Alongside with the use of the sun and star compass [39], the use of the Earth's magnetic field in long-distance migration is at present a leading hypothesis on how migratory birds orient which is therefore also this sub-chapter's focus. Further theories are discussed below.

In 1972, Wiltschko & Wiltschko found that European robins use the inclination (the angle between the magnetic field vector and the horizontal) of the Earth's magnetic field lines rather than the polarity (like a conventional man-made compass) of the Earth's magnetic field for navigation [40], opening the question of how long-distance migrators, who cross the equator where the inclination angle is 0°, might continue navigating on the other side where the inclination angle has the opposite sign. The question was answered by the authors in 1992 who found that garden warblers reversed their response to the Earth's magnetic field inclination after having been exposed to a simulated 0° [35]. It was later found that songbirds might not only use the inclination, but also the magnetic declination (the difference

between geographic and magnetic North or South) which expands their compass sense to a bicoordinate map. However, not all bird species seem to be equally capable of using this type of map since Eurasian reed warblers were found to use declination information whereas the European robin and garden warblers tested under similar conditions were not able to do so [37, 41]. In addition to the question of how migratory birds can sense in which direction they need to fly, another interesting question is how they know when they have arrived. Wynn *et al.* found that Eurasian reed warblers solve this by remembering the inclination present at their natal or breeding site during the previous year before returning. However, they also found this to be the only magnetic cue the birds seem to use, contradicting the idea of a bicoordinate map [38].

Despite extensive tracking capabilities there are still a lot of open questions with respect to avian navigation at the behavioural level, partly due to contradicting findings (for example if and how a bicoordinate map might be used) or by new questions opening up due to experimental evidence such as the finding that some birds are able to completely reverse their migration direction within a time frame as short as two years [42], which raises the question why they might have evolved this capability.

Directional orientation in migratory birds has been found to depend on either white or low-level monochromatic light ranging from UV to green light [43] with several bird species becoming inactive or disoriented under overcast skies [2, 44]. It has further been suggested that the magnetoreception processes might take place in the eyes [5] which was later confirmed experimentally including the additional finding that the right eye of several bird species is responsible for navigation, not the left [45, 46]. However, this claim has since been contested [47–49].

The magnetic sense used for navigational purposes was first described for the European robin mentioned above and was subsequently shown for further night-migratory passerines [35]. Interestingly, for the European robin, amongst other birds, it has been found that migratory behaviour is genetically determined [36]. In subsequent years a lot of research was focused on this particular night-migratory bird.

To name a few of these findings, European robins are able to navigate under blue light conditions, but not under yellow or red light [50], together with garden warblers they have been found not to respond to magnetic declination [41], and their retinas have been the subject of intense research in order to identify the protein that is suspected to be involved in magnetoreception as well as its whereabouts [51–54]. It has also been reported that European robins can orient under green light conditions [50].

1.4.1 Different theories of avian navigation

Over time, several theories of avian navigation have been discussed with some studies pointing out that different compass strategies might be used by different bird populations. Compasses using celestial cues such as the stars, sunlight, or skylight polarization patterns have been discussed [39] as well as olfactory cues [4]. For example, for Cory's shearwaters it has been shown that they are disoriented in the absence of a functional olfactory system but still capable of navigating large distances across oceans when they were magnetically disturbed by carrying around a freely tumbling magnet in a box on top of their heads while having an intact olfactory system [55]. However, the two leading hypothesis for long-distance avian migration are the so-called radical pair mechanism, based on the perception of the inclination of the Earth's magnetic field lines, on which this thesis will focus and which is discussed in detail in section 1.5 as well as a theory based on magnetite which consists of iron-containing nanoparticles [2] and is further discussed below.

Using the principle of electromagnetic induction, where an electrical conductor moves through a static magnetic field, some elasmobranchs such as sharks and skates can detect weak electric fields when moving through the Earth's magnetic field which induces a detectable voltage change in their bodies. They are thought to detect this via their ampullae of Lorenzini which are sensitive to voltage-gradient changes in the surrounding medium [17, 56]. Several theories have evolved around how these animals might be able to use these capabilities as a compass or to detect prey [57, 58]. In 2019, it was proposed by Nimpf *et al.* that pigeons might be able to navigate using the same principle of electromagnetic induction. They were

found to have a comparable mechanism to sharks and skates capable of detecting changes in electric fields expressed in the semicircular canals of their inner ear and dependent on voltage-gated calcium channels found on their electrosensory hair cells [59]. This work was presented as exploring a theory that had been forgotten about without having been proven or falsified. The same research group has also been involved in the exploration of pigeon light-dependent magnetoreception via cryptochromes as part of the work on magnetoreception via the radical pair mechanism [60], illustrating the breadth of possible explanations for the various open questions within this field of research.

Another theory of avian navigation, originally thought to be a competitor of the radical pair mechanism but nowadays, depending on the species, seen as either potentially complementing it or not relevant, is magnetite-based magnetoreception: crystals of single-domain magnetite (Fe_3O_4) are the ferrimagnetic material thought to be at the heart of this theory. Their direction of magnetisation could be changed by strong magnetic fields. Magnetotactic microbes, bacteria and eukaryotic algae, have been found to contain chains of magnetite particles known as magnetosomes. These chains have a magnetic moment large enough to be influenced by an external magnetic field that can rotate the organisms into alignment with this field, making it a passive rather than an active process of magnetic field perception [3].

Magnetite was also suggested to be found in the beak of pigeons and in several fish as these can be conditioned to respond to strong magnetic fields. Indeed, clusters of small crystals were found in the skin of pigeons' upper beaks [11, 61]. However, it was subsequently shown that the suggested magnetite-containing cells are most likely macrophages which are not connected to the birds' sensory system. However, the possibility of the involvement of magnetite in magnetoreception could not be fully excluded [62, 63].

There is some evidence that the trigeminal nerve, which enters the upper beak, is involved in magnetoreception. Since magnetite seems to occur in the beak of certain species, it was suggested to be involved in magnetoreception. However, since the trigeminal nerves are not necessary for magnetic compass orientation, it was

suggested that magnetite might be involved in some sort of map-sense rather than the perception of the direction of the Earth's magnetic field [2].

Recently, another theory has emerged, called the level crossing mechanism, based on rotating macromolecules as the magnetosensors [64]. However, this theory is at present more of a suggestion which is not backed up by sufficient experimental results as none of the putative magnetosensors or any interaction partners have been identified [65].

1.5 The radical pair mechanism

The radical pair mechanism of magnetoreception is a well-established theory explaining how the Earth's magnetic field might influence a photochemical reaction involving radical pairs in a way that could encode directional information necessary for avian navigation [2].

The general idea of the radical pair mechanism is displayed in Figure 1.2. A and B, two reactants, are exposed to light which excites reactant A. Subsequent electron transfer takes place with B as the donor and A as the acceptor. As the reactants are usually initially in a singlet state with all electrons paired up and as the formation of radical pairs is spin selective, following electron transfer the radical pair exists in its singlet state. The required spin conservation arises from the requirement to conserve the total angular momentum. The two radicals so formed experience singlet-triplet interconversion which depends on the strengths of the hyperfine interactions between the unpaired electrons and spins of surrounding nuclei. Local magnetic field differences caused by the different hyperfine states of the individual radicals (or sometimes differences in g-values if they are sufficiently large) allow this interconversion [66]. The interconversion and therefore the yields of singlet and triplet products can be modulated by an externally applied magnetic field as the spin magnetic moment of the electrons interacts with the magnetic field. Since the hyperfine interactions are anisotropic, any changes in the magnetic field orientation influence the singlet-triplet interconversion. As only the radicals in the singlet state can undergo the reverse reaction back to the ground state due to

the required spin conservation, indicated by k_{r}, the magnetic field would influence the probability of the reverse reaction [2, 25]. Both the singlet and triplet states of the radical pair can lead to product formation, indicated by k_{sf} for the singlet state and k_{tf} for the triplet state. It is possible, that these reactions occur at different rates and yield different products depending on the radical pair's initial state. Usually, interactions with the Earth's magnetic field would not be considered capable of influencing a chemical reaction since their energies are about a million times smaller than $k_{\mathrm{B}}T$, where k_{B} is the Boltzmann constant, the energy associated with random Brownian motions, so therefore these interactions have a negligible effect on thermal rates of reaction. However, as the above described radical pair state is a system highly out of equilibrium immediately after its creation, even a very small influence such as the Earth's magnetic field can tip the resulting product yield in one direction or another [2]. Aside from in the context of magnetoreception, the radical pair mechanism has been observed in other effects, such as the chemically induced dynamic electron polarization and the chemically induced dynamic nuclear polarisation, which were both found as early as 1967, as well as the magnetic isotope effect where isotopic substitution was found to have an effect on a reaction within the radical pair mechanism framework [66]. However, these varieties will not be discussed further in this chapter as the focus of this work lies on magnetoreception.

In 1978, Klaus Schulten and Charles Swenberg proposed a mechanism involving radical pairs whose singlet-triplet interconversion could be influenced by the Earth's magnetic field depending on its orientation with respect to the reactants, therefore enabling an animal to navigate using magnetoreception [67]. In 2000, it was argued theoretically that magnetic fields as weak as the Earth's and weaker could influence the effect of the anisotropic hyperfine coupling of unpaired electrons and nuclear spins leading to different reaction yields depending on whether the singlet or triplet state was favoured. An additional requirement for this theory is that the radicals would have to be permanently aligned within the organism as due to the anisotropy of the hyperfine interactions, the orientation of the radicals with respect to the Earth's magnetic field lines is essential [5]. The same work published

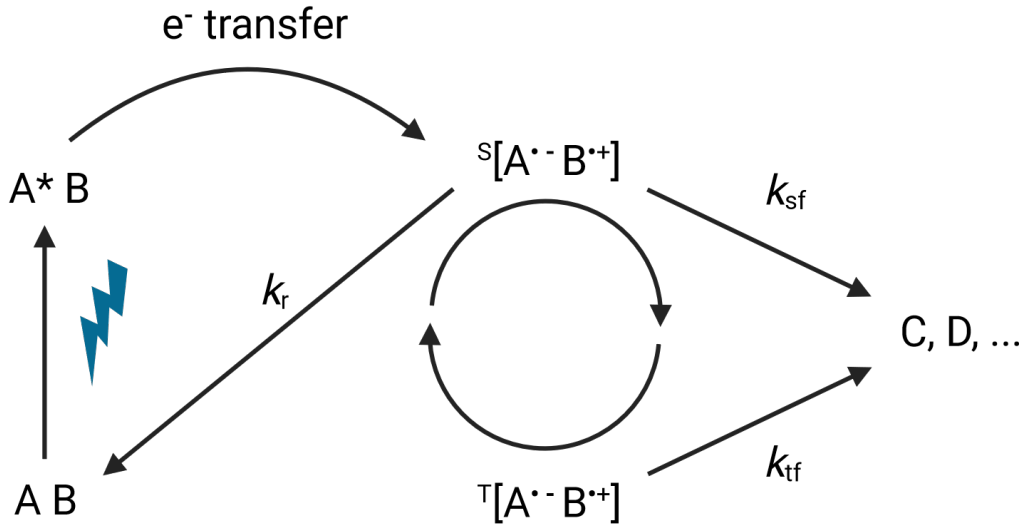

Figure 1.2: Simple reaction scheme of the radical pair mechanism. Reaction scheme of a magnetic-field-dependent radical pair interconversion between singlet and triplet states. The initial step is a photoexcitation of reactant A leading to a subsequent electron transfer from reactant B to reactant A. Figure modelled after Ritz, Adem and Schulten (2000) [5] and Hore and Mouritsen (2016) [2]. Created with BioRender.com.

by Ritz *et al.* also proposed cryptochromes as the putative magnetoreceptors which are further discussed below.

Cryptochromes are a protein family closely related to photolyases but lacking their DNA-repairing capabilities. They non-covalently bind a flavin adenine dinucleotide (FAD) molecule as their cofactor and comprise a photolyase-homology region (PHR) as well as a variable C-terminal extension that distinguishes them from photolyases [68].

In 2013, Liu *et al.* calculated electron transfer rates within a photolyase. This in combination with femtosecond spectroscopy showed that an electron transfer chain comprising a triad of tryptophans displays a reduction-potential gradient favouring the direction towards the FAD with <150 ps electron transfer to the FAD and much slower (nanosecond) electron transfer in the opposite direction along the Trp chain [69]. These findings speak in favour of a unidirectional electron transfer flow enabling the reduction of FAD and the formation of a radical pair distanced in space comprised of the cofactor and the terminal member of the Trp chain. This

has later also been repeatedly found in cryptochromes [2, 70, 71].

For a while, the Trp chain was thought to consist of three members, which was later extended for some cryptochromes to a fourth Trp and/or a nearby tyrosine at the end of the electron transfer chain, for example when the crystal structure of pigeon (*Columba livia*) cryptochrome 4, *Cl*Cry4a, was solved [72] where a Trp tetrad was found. The cryptochrome found in *Drosophila melanogaster* (*Dm*Cry), the fruit fly, was also found to contain a fourth Trp and *Cra*Cry, a cryptochrome from the green alga *Chlamydomonas reinhardtii*, has a Trp triad and a tyrosine in the position of a fourth aromatic amino acid [73–75].

The theory of the radical pair mechanism, often abbreviated as RPM, is supported by numerous studies showing that magnetic fields within a 1-100 mT range have an effect on radical pairs. However, the theory has not unequivocally been shown to be responsible for animal magnetoreception [2], but over the years numerous studies have been conducted on this subject amassing a wealth of information on the general subject of magnetoreception [2, 17, 25, 76, 77].

Figure 1.3 displays the current understanding of the RPM within cryptochromes in some more detail. In (A) the electron transfer steps are exemplified within the European robin cryptochrome 4a and (B) displays the actual mechanism. FAD gets photoexcited via blue light in its singlet state, followed by subsequent electron transfer steps via the nearby tetrad of electron donating aromatic amino acids. The singlet-born radical pair can subsequently either fall back into its ground state or undergo singlet-triplet interconversion which can be influenced by an external magnetic field. Upon deprotonation of the Trp cation radical, a FAD anion radical and a neutral Trp radical could form a stabilised radical pair, followed by a protonation step of the FAD radical and oxidation of the neutral Trp radical leading to a Trp in its ground state and a long-lived neutral FAD radical as the signalling state.

Strong support for the radical pair mechanism specifically in birds comes from several studies which show that radiofrequency magnetic fields cause disorientation in several migratory birds. For example, Ritz *et al.* applied oscillating magnetic

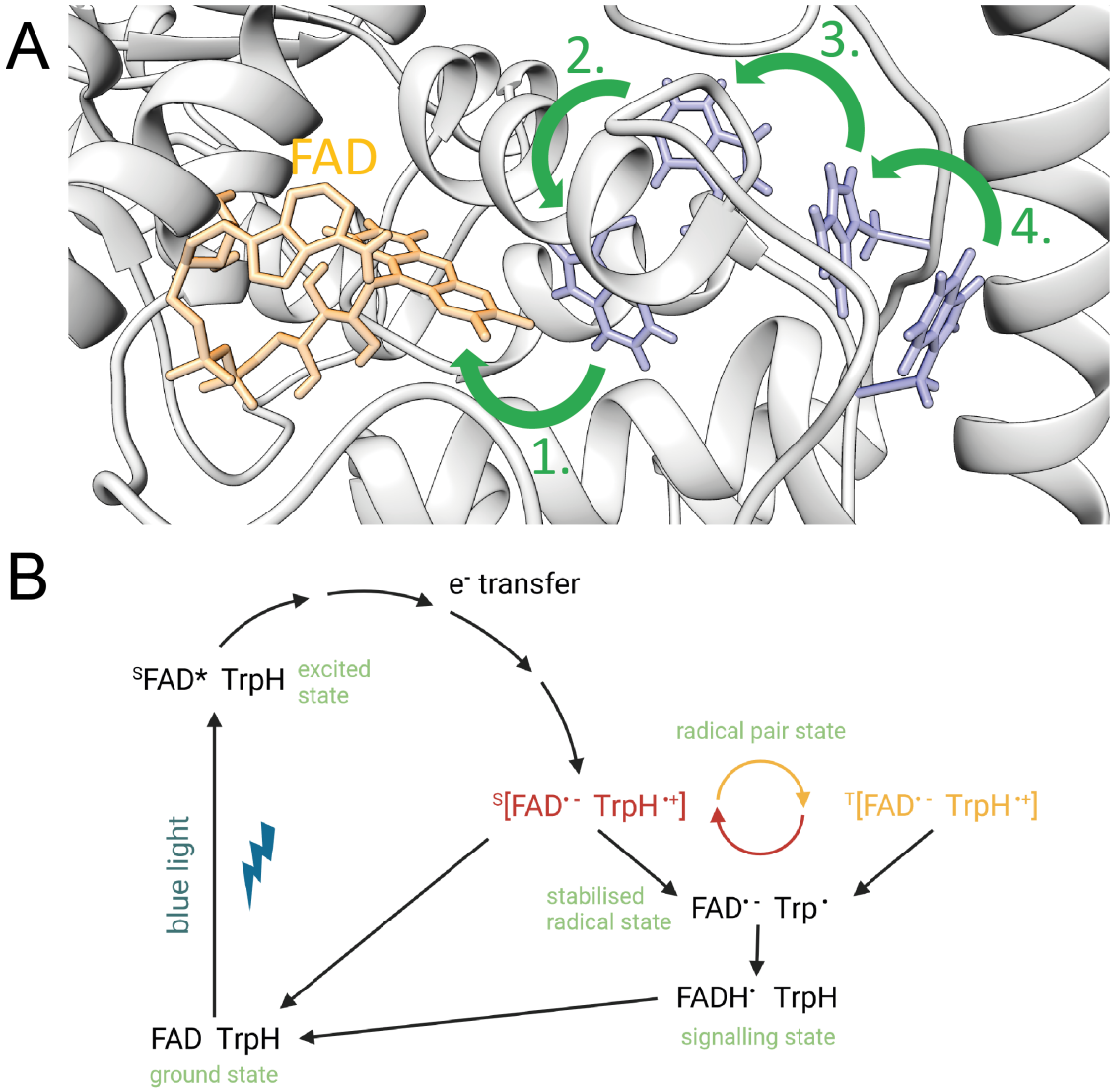

Figure 1.3: Proposed photochemical scheme for photo-activation and and radical pair formation in cryptochromes. (**A**) Electron transfer steps of cryptochromes illustrated using a European robin cryptochrome 4a homology model published elsewhere [78]. (**B**) The interconversion between singlet and triplet radical pairs is in a non-equilibrium state that can be influenced by the Earth's magnetic field to impact on the yields of the stabilised radical state and signalling state. Reaction scheme modelled after Xu *et al.* [71]. Created with BioRender.com.

fields at an angle to the Earth's magnetic field which disoriented European robins [79] and Engels *et al.* showed that the same bird species was disoriented just by exposure to anthropogenic electromagnetic noise until it was attenuated by about two orders of magnitude [80]. According to quantum chemical calculations [81] suggesting subsequent behavioural experiments performed by Leberecht *et al.*, electromagnetic noise in the frequency range of 75-85 MHz disrupts the orientation of a night-migratory songbird [82]. This finding was further analysed and led to migratory birds having an upper limit of the radiofrequency induced disorientation between about 80-145 MHz. This is consistent with a magnetoreceptor based on a flavin-containing radical pair rather than magnetite as in this case radiofrequency fields would not be expected to disrupt the magnetic orientation [83].

In 2009, Ritz *et al.* subjected migratory birds to magnetic fields that oscillated at half, double and exactly the Larmor frequency, the precession frequency at which an electron spin would precess around the axis of a magnetic field. If a sufficiently long-lived radical pair is present within the birds, a magnetic field oscillating at the Larmor frequency should be capable of interfering with the radicals' precession considering the mathematical relationship of $\nu_{\mathrm{L}} = \gamma_e B_0/2\pi$ where ν_{L} is the Larmor frequency, γ_e the gyromagnetic ratio of the electron, and B_0 the intensity of the magnetic field. As a result, disorientation of the bird would be expected (given one of the radicals has no or extremely small hyperfine interactions). Indeed, at low magnetic field strengths (15-150 nT) the magnetic field oscillating at the Lamor frequency disoriented the tested birds [84]. However, this study could not be replicated independently in a later study pointing out the difficulty to understand and measure Larmor frequency effects in a noisy biological environment [85]. Instead, it was found that weak broadband electromagnetic oscillating fields can disrupt compass orientation [80, 85], a finding that as of yet still lacks an explanation.

There are still many discussions in the literature on how to relate the theoretical calculations and the RPM in general to the behavioural results which often seems difficult to do [81]. However, there are several approaches to try and tackle this problem, such as theoretical calculations finding that a radical pair consisting of a

FAD radical and a radical $Z^{\bullet}$ that has no or extremely small hyperfine interactions could be two orders more sensitive than the conventional FAD and Trp radical pair (if the electron-electron dipolar coupling is very small) [86]. This idea was later extended to include a third, so-called 'scavenger' radical enabling the prediction of larger magnetic field effects which could operate under noisy, biological conditions [87]. However, these scavengers have yet to be identified and experiments working with Earth's strength magnetic fields and cryptochromes showing significant magnetic field induced effects, which would be a milestone towards proving that the RPM is an important component for avian navigation, still have to be conducted.

1.6 Cryptochromes

Cryptochromes are 60-70 kDa proteins found in all biological kingdoms[76, 77, 88–91] which consist of a conserved PHR and an intrinsically disordered C-terminal tail (CTT) [92] of varying lengths between ∼25-150 amino acids [91]. Their PHR domains share structural homology with the DNA photolyases, enzymes which repair UV-B-damaged DNA using blue and UV-A light [76, 90, 93]. Many cryptochromes bind the chromophore flavin adenine dinucleotide (FAD) as their cofactor, embedded in a FAD binding pocket on the inside of the PHR [94]. Some members of the cryptochrome/photolyase family also bind a second nucleotide-like cofactor within the PHR. Methenyltetrahydrofolate (MTHF) is an example of such a light-absorbing chromophore that has been found in cryptochromes. However, due to low sequence conservation of the MTHF binding site and its apparent non-specificity, theoretically allowing other nucleotide-like molecules to also bind in its position [76], most research focuses on the FAD cofactor. The PHR provides the scaffold to which all known cofactors bind and varies slightly between different proteins. For example, photolyases have a small FAD-access cavity at their catalytic site whereas *Arabidopsis thaliana* Cry1 displays a larger cavity in the same region [95].

Cryptochromes first appeared as the 'cryptic' photoreceptors in cryptogams, a term for certain types of plants that nowadays is obsolete [98]. Previously, they were often referred to as 'blue/UV-A light photoreceptors', referring to a

Table 1.1: Different types of cryptochromes. The three different classification systems (1), (2), and (3) were used over time to classify cryptochromes. They have been used somewhat interchangeably and not necessarily linearly in time. They are each divided into different types. The most recent and most elaborate classification discussed in the main text is the one detailed in (3). This table is not intended to be an exhaustive list, but acts as a summary and overview of the most important classification systems mentioned in the main text which regularly appear in the relevant literature. However, for an insight into an updated and more in-depth classification system, the reader is referred to a recent publication by Ozturk [96].

(1)		**(2)**	**(3)**		
Type	**Details**	**Type**	**Type**		**Details**
plant	plants	plant	plant		
Drosophila-like type 1	invertebrates, photoreceptors sensitive to blue/UV-A light [97]	animal	animal	I	light-dependent, FAD, *Drosophila*-like type
				II	light-independent, no FAD, mammalian-like type
Vertebrate-like type 2	vertebrates, invertebrates, involved in transcriptional feedback loop of the circadian clock [32]			III	redundant
				IV	light-dependent, FAD
		Cry-DASH	single-stranded DNA-specific/Class 0 photolyase		

photolyase without the DNA repair activity [99]. They were found to be responsible for hypocotyl growth regulation in *Arabidopsis thaliana* after the HY4 gene was isolated and found to encode a protein that binds FAD and lacks the DNA-repairing mechanism that defines photolyases [100, 101], subsequently named Cryptochrome 1 [102]. Another *Arabidopsis thaliana* cryptochrome, Cry2, was found soon afterwards and shown to regulate floral initiation [103]. The next plant cryptochrome, Cry3, was named a Cry-DASH protein [104], a protein that retained the DNA or RNA binding capability typical for photolyases. Once cryptochromes had been discovered in plants, other organisms were successfully probed for these proteins. These widely present proteins were found in organisms ranging from bacteria to fruit flies, mice, and humans and initially divided into *Drosophila*-like type 1 Crys (only found in invertebrates) and vertebrate-like type 2 Crys (found both in invertebrates and vertebrates) as well as three different classes: plant cryptochromes, animal cryptochromes, and CRY-DASH proteins [95]. Nowadays, after years of research on various different types of cryptochromes, the classification system has slightly changed to dividing the proteins into plant cryptochromes and animal cryptochromes which are further divided into *Drosophila*-like type I Crys, mammalian-like type II Crys, and type IV Crys (Cry4) which have so far been found in birds, amphibians, fish, and reptiles. Types I and IV are light-dependent and contain a FAD cofactor. Type II Crys are the opposite, they are light-independent and lack the FAD [72, 105]. Initially, one of the cryptochromes found in zebrafish was classified as a type III Cry, but it was found that it actually belongs to the type II Crys, rendering type III Crys a redundant classification [90]. The Cry-DASH proteins have recently been named single-stranded DNA-specific photolyases or Class 0 photolyases, being more closely related to the photolyases and plant cryptochromes than any of the other cryptochrome members of the family [96]. For a better overview, the classification systems that evolved over time are listed in Table 1.1. A few years ago, a Cry5 variant appeared [106]. However, so far it does not seem to have made it into the common classification systems and therefore will not be further discussed here. Additionally, multiple splice isoforms were found of Cry I, Cry II and Cry IV

in different avian species such as in the European robin (Cry1a, Cry1b, Cry2, Cry4a, Cry4b) [51, 53], and the pigeon (Cry2a and Cry2b) [107]. The European robin's Cry4a has been shown to have several of the properties required for a magnetoreceptor. It is stably expressed throughout the day whereas the expression pattern of its splice variant Cry4b oscillates diurnally, suggesting that the two variants might have different functions [53].

Several types of cryptochromes have been found in birds' retinas. They were first found in chicken retinas and subsequently in retinas of the European robin. Cry1b has also been found to be expressed in a seasonal pattern in retinal ganglion cells of the European robin [108] and Cry2 was observed in a broad range of locations within the retina in European robins, homing pigeons and chickens, pointing towards its role as a transcription factor rather than being a part of the magnetoreception molecular machinery [54]. However, despite Cry1 and Cry2 seemingly fulfilling different roles other than being involved in magnetoreception, Cry4 has been found to be consistent with the proposition to be the magnetoreceptor as it is localised in the double cone photoreceptor cells within the retina, binds FAD, and follows a seasonal expression pattern in migratory Europeans robins, but not in non-migratory chickens. The finding that Cry4s are found in double cones, or more specifically in the outer segments of double cones, is important as these areas contain hundreds of parallel cell membranes that could facilitate alignment of these proteins [52] which is a necessary requirement for them to function as a magnetic direction sensor (see section 1.6.1). The above mentioned retinal ganglion cells can be found in the eyes (where they are in fact the one type of cells all information leaving the eye has to travel through) as well as a part of the brain called Cluster N, one of the most active parts of the brain during the processing of magnetic-compass orientation which also processes visual information [2]. However, a direct connection between a specific cryptochrome and the transmission of magnetic-compass information through these cells has not been established.

According to a sequence analysis of plant and animal cryptochromes, it appears to be the case that they developed independently from each other. Mammalian and

fly cryptochromes are less closely related to plant cryptochromes than they are to (6-4)-photolyases. However, as they have the same or at least very similar functions, the genetic basis for this seems to have developed separately in both cases, which is also called repeated evolution [109]. From this follows that the exact mechanism of function might be different in plant and animal cryptochromes including their likelihood to form dimers or higher order oligomers as discussed below.

1.6.1 Cryptochrome's involvement in the RPM

In addition to the studies on plant cryptochromes mentioned above, a lot of studies on *Dm*Cry), often used as a representative of animal cryptochromes, have been published over the years [25], one example being the work of Yoshii *et al.* who showed that *Dm*Cry mediates the circadian clock in a magnetosensitive way [110]. It was also found that the *Drosophila* cryptochrome could be replaced with a cryptochrome from another organism, the monarch butterfly, which restored the fruit flies' light-dependent magnetic sense [97]. Cryptochrome research in mammals ensued soon after the first cryptochrome had been identified. Two human flavoproteins with a high sequence similarity to *Drosophila melanogaster* (6-4) photolyase (these proteins recognise pyrimidine-pyrimidone (6-4) photoproducts which accumulate due to UV-exposure [93]) were identified as cryptochromes in 1996 [111] and found to act as components of the circadian clock, but in this case in a light-independent manner [112]. The possibility of the involvement of mammalian cryptochromes in magnetoreception is still discussed and only very recently have a few more hypotheses about their possible involvement been put forward that still need to be experimentally investigated [105]. However, since most mammals seem to only express Type II cryptochrome which has a low binding affinity for FAD, this protein has no photosensitivity and therefore the formation of a radical pair seems unlikely. However, it would be conceivable that the proteins could bind FAD *in vivo* or would otherwise be involved in magnetoreception without the necessity to contain a radical pair [105, 113].

So far, cryptochromes are the only vertebrate proteins known to form radical pairs following photoexcitation [82], making them a strong candidate for a protein perceiving the Earth's magnetic field via a radical pair. Recently, European robin cryptochrome 4a has been shown to be magnetically sensitive *in vitro* and to form the radical pair [$FAD^{\bullet-}$ $TrpH^{\bullet+}$] upon photoexcitation [71].

Whichever protein might fulfil the function of a magnetoreceptor, due to the involvement of hyperfine interactions at the core of the RPM, which are generally anisotropic, a defined directionality of the proteins is required. This would be facilitated by a strict orientation of the magnetoreceptor within the organic tissue of the navigating species, e.g. via membrane association. However, cryptochromes are soluble proteins that do not directly associate with a membrane. This puzzle could be solved if they were found to associate with membrane proteins [2]. *Dm*Cry was indeed found to be linked to membrane depolarization and elicit an increased neuronal firing rate via the voltage-gated potassium channel β-subunit Hyperkinetic after light-activation [114]. Recently, six putative interaction partners of cryptochrome 4a were identified amongst which were a voltage-gated potassium channel and the α-subunit of a cone-specific heterotrimeric G protein called guanine nucleotide-binding protein G(t) subunit α-2 (GNAT2) which are two examples of proteins that could facilitate the required membrane-association [115, 116]. Another theory was published by Qin *et al.* in 2016 which explores the idea that the iron-sulphur protein IscA1, also called MagR, might interact with cryptochrome to form a complex that could act similarly to a compass needle. However, several studies published later contested the validity of this theory [115, 117].

1.6.2 Cryptochrome's C-terminal tail

The C-terminal tail (CTT) of cryptochromes has been suggested and found to be responsible for downstream signalling multiple times. In the context of modulating circadian rhythms, it was found that removal or modification of the CTT leads to altered periodicity of the circadian rhythm thereby attributing to the CTT the ability to fine-tune Cry functionality [92]. For example, a truncated *Dm*Cry missing

its CTT interacts light-independently with a protein called TIM and promotes its degradation, whereas the presence of the CTT makes this behaviour light-dependent [118]. Similarly, studies investigating *Arabidopsis thaliana* Crys (*At*Crys) and a protein called COP1 also found the C-terminus to be of major importance for the light-dependent interaction regulation [119–121]. In *At*Cry1 dimerisation happens via the N-terminal domain and is required for light-activation of the C-terminal domain [122] and in *At*Cry2 the C-terminus is involved in light-induced oligomerisation. This process can even be enhanced by employing C-terminal changes such as different types of C-terminal extensions [123]. The chicken retinal Cry4's C-terminus was found be loose and accessible for protein-antibody interactions in the dark, but closely attached to the protein and not available for interaction upon exposure to blue light [124]. Another interesting task worth mentioning that the CTT has been found to fulfil is the influence on the redox state of FAD in a plant-like cryptochrome of diatom microalgae: in the absence of the CTT, FAD is in its oxidised state whereas in proteins with their CTT attached FAD was found to be present as a neutral radical [125].

However, the structure and exact position of the C-terminal tail are generally quite elusive and up until 2011 the only two solved cryptochrome crystal structures did not contain the C-terminal tail: the cyanobacteria *Synechocystis* Cry does not have a C-terminal extension and for *At*Cry1 the C-terminus had to be truncated to permit crystallisation. In 2011, the crystal structure of full-length *Dm*Cry was solved [126, 127]. Since then further cryptochrome structures have been solved, such as the PHR of mouse Cry1 [128] and the pigeon Cry4, the first crystal structure of a vertebrate Cry with functional photochemistry [72]. However, up until today, there is no second structure solved which includes the full C-terminal tail.

A great deal of work has been done using *Dm*Cry given that it is the only cryptochrome of which a full-length structure has been solved. In 2013, Vaidya *et al.* found that photoreduction of *Dm*Cry led to the release of the CTT [129]. This was backed up by a study performed by Chandrasekaran *et al.* using pulsed dipolar electron-spin resonance spectroscopy which found that flavin photoreduction shifts

the CTT by about 1 nm, but only if the cofactor was present in its charged radical state [130]. Deletion of the CTT in fruit flies disrupted magnetic field responses in another study where negative geotaxis (climbing upwards against gravity) was monitored and found to be both Cry- and light-dependent. The flies with the truncated cryptochrome were constitutively climbing upwards, light-independently. The same study also found that the mutation of the terminal member of the Trp triad (the study was published in 2014, before the fourth member of the Trp tetrad had been identified [74]) does not abolish the responsiveness of flies performing magnetic field-dependent negative geotaxis [131], possibly indicating that the third member of the Trp triad might not be necessary to induce conformational changes ultimately resulting in behavioural responses. Gegear *et al.* found in 2010 that they could recover a magnetic response in *Drosophila melanogaster* using monarch butterfly cryptochromes that are not responsive to wavelengths above 420 nm, which would point towards a magnetoreception mechanism that might not need the FAD cofactor to be photo-excited [97]. However, more recently, another study by Bassetto *et al.* was published which could not replicate previous studies performed by Gegear *et al.* [97, 132], Foley *et al.* [32] and Fedele *et al.* [131] and the authors concluded that they could find no evidence for the reported magnetically affected behaviours in *Drosophila*, neither when investigating their ability to associate sucrose with a magnetic field nor when looking at their apparent negative geotaxis [133].

In another study by Fedele *et al.*, it was demonstrated that the *Dm*Cry CTT can have an impact on electromagnetic field-induced circadian period changes [134]. A recently published work by Bradlaugh *et al.* follows up on this finding and shows that just FAD and the C-terminal tail are sufficient to elicit a magnetic response in *Drosophila* larvae cells with increasing concentrations of FAD augmenting the magnetic field and blue light effects mediated by the CTT. The authors conclude that parts of cryptochrome might be essential for magnetoreception, but not the whole protein as long as the essential components are present [113]. Several interesting motifs alongside the CTT of *Dm*Cry have been found and partially experimentally verified to be important for downsteam interactions with other proteins, an example

being a calmodulin binding domain which binds calmodulin in a calcium-dependent manner which is most likely involved in visual transduction [135, 136].

Much less work on elucidating the function of avian CTTs has been published so far. Möller *et al.* found that the CTT of European robin Cry2 contains a nuclear localisation signal making it less likely for this protein to be involved in magnetoreception [51]. Several computational studies have been performed on European robin and pigeon Cry4a showing the release of the CTT in the case of the migratory European robin but not in the non-migratory pigeon [137, 138]. However, there is still a lot of work to be done on avian cryptochromes in the next few years.

The above described findings lay bare how many different functions the different types of cryptochromes and specifically their CTTs might fulfil and they point out that a lot of questions regarding the how and why are still left unanswered.

1.6.3 Cryptochrome oligomerisation

A lot of proteins have particular functions in their oligomeric states such as different modes of action or a changed structural stability compared to their monomeric forms [139, 140].

Some proteins of the cryptochrome family have been found to fall into that category. The plant cryptochrome *At*Cry2 was found to form a tetrameric state under full spectrum light during crystallisation which was indicated to be the result of light-induced oligomerisation events [141]. *At*Cry1 as well as maize Cry1 were also found to form dimers and tetramers [122, 142]. However, it should be noted that the conformational changes observed for plant proteins that lead to their oligomerisation have so far not been observed in mammalian cryptochromes. Recently, L-Cry from the bristle worm *Platynereis dumerilii* has been proposed to form homodimers in which the monomer units have different light sensitivities, enabling the proteins to act as light-sensors that can differentiate between different light intensities and durations [143].

Interestingly, the animal-like algal cryptochrome *Cra*Cry was seen in its dimeric form *in vitro* when running the protein through a size-exclusion chromatography

column in the dark. When the CTT was removed and the protein subjected to blue and then red light (to convert FAD_{ox} to $FADH^{\bullet}$ with blue light and $FADH^{\bullet}$ to $FADH^{-}$ with red light) this facilitated disulphide-bonded oligomerisation [144].

1.7 Structural investigation of light-sensitive proteins using HDX-MS

Multiple light-sensitive proteins or photoreceptors have been studied over time using the technique of hydrogen/deuterium exchange mass spectrometry (HDX-MS) which is especially powerful at elucidating dynamic changes over time in solution at ambient temperatures for a wide range of molecular sizes. It is often used to complement findings obtained by other techniques such as X-ray crystallography, small- or wide-angle X-ray scattering, nuclear magnetic resonance (NMR), or cryo-electron microscopy. In 1999, the first photoactive protein analysed via HDX-MS was the photoactive yellow protein from the *p*-hydroxycinnamic acid binding xanthopsin protein family, one of the six big cofactor-binding photoreceptor families (the other ones being retinal binding rhodopsin, bilin binding phytochromes, and the flavin-containing photoreceptors which are subdivided into blue-light-using-FAD (BLUF) domain containing, light-oxygen-voltage (LOV) sensing proteins, and cryptochromes) [145]. Substantial conformational changes were reported for this protein upon light-activation [146, 147]. HDX-NMR was also applied to the same protein, revealing changes near the cofactor [148]. However, the latter findings did not include the more extensive conformational changes that could be captured using HDX-MS. An important strength of HDX-MS is exactly this, the possibility to monitor long-range conformational changes, often initiated by a light-sensitive cofactor, which, together with other techniques that look in more detail at specific parts of the proteins, can yield a wealth of information. Using a mix of different techniques and HDX-MS, homodimerisation upon light-excitation, undocking of a single protein domain, hydrogen-bonding network disruptions or rearrangements, or transient unfolding can be elucidated. For example, LOV and BLUF domains

have been found to include a C-terminal region packed against their core which seems to be involved in signal transduction [145].

To name a few more examples, Partch *et al.* demonstrated in 2005 for the first time light-dependent conformational changes in the C-terminal region of a cryptochrome species, working with *At*Cry1 [118]. A few years later, in 2016, Von Horsten *et al.* investigated phytochrome B, a photoreceptor similar to cryptochromes with the covalently bound, light-responsive cofactor phycocyanobilin, also using HDX-MS. They found that light-induced conformational changes involve complex rearrangements of different protein regions which ultimately lead to altered downstream interactions with other proteins [149]. This finding is interesting as it illustrates how concerted motion might be the response to photo-excitation of a chromophore rather than the movement of a singular protein region. It further shows the capabilities of HDX-MS to analyse conformational changes that were elicited due to different light conditions and which will ultimately affect downstream signalling.

Winkler *et al.* showed that it is also possible to analyse light-responsive complexes of proteins and that HDX-MS is not limited to single proteins [150]. Over the years, HDX-MS has been optimised both in terms of workflow and data analysis and has also seen a partial workflow automation. It is now a valuable tool even for complicated systems such as the analysis of membrane-embedded proteins [151, 152] including for example bacteriorhodopsin, a light-driven proton pump which proved to be challenging to be analysed in its naturally occurring trimeric state, but could be investigated as a monomer in bicelles [153].

Native mass spectrometry, a method capable of measuring whole proteins in their native or near-native state, has also already been employed to investigate photoreceptor function, for example by Camacho *et al.* who probed the conformational landscape of the UVR8 photoreceptor [154] followed by an investigation of CarH, a B_{12}-binding protein, by the same group [155]. Another example of using this technique for photoreceptors was published by Chen *et al.* who investigated the signalling cascade of rhodopsin [156]. In both cases a modified native mass spectrometer which subjected the proteins to light just before measurement was

used. However, the focus of this work is centred on an investigation using HDX-MS with native mass spectrometry supplying additional supportive data.

1.8 Contribution

Magnetoreception, the capability of living organisms to detect the Earth's magnetic field, has been researched for over 50 years. Early on, in 1972, it was shown that migratory birds, more specifically European robins, use the inclination of the geomagnetic field for orientation rather than the polarity which differentiates their sensory mechanism from human-made compasses [40, 157]. Since then the subject has amassed a wealth of studies distributed across a wide variety of research fields. Behavioural experiments, molecular biology approaches, biophysical investigations and theoretical calculations are among the methods of choice of this highly interdisciplinary field in which the European robin remains one of the most researched species capable of navigating long distances [2].

This thesis aims to analyse the photoreceptor Cry4a, the putative magnetoreceptor in birds which has been shown to have the capability of responding to the inclination of the Earth's magnetic field [52, 71], with a focus on using the technique of HDX-MS. Further methods that are introduced and which were used to complement the HDX-MS data are native mass spectrometry (MS), mass photometry (MP), chemical crosslinking followed by mass spectrometric analysis (XL-MS), electron paramagnetic resonance (EPR) spectroscopy, and molecular dynamics (MD) simulations, all of which are introduced in chapter 2.

Chapter 3 introduces the first experimental results focusing on European robin Cry4a and one of its functional mutants. It takes the reader step by step through different HDX-MS approaches and difficulties that were encountered during the work on these proteins. It also focusses on the C-terminal tail of the protein and the difficulties in terms of data analysis that come with this particular experimental approach despite reliably yielding data on how surface-accessible every part of the protein is under different light conditions.

Chapter 4 introduces successful repeats of the experiments described in Chapter 3 on European robin Cry4a. Additionally, it introduces two further cryptochrome 4a species, one from the migratory quail and one from the non-migratory chicken. It analyses in depth the changes in surface-accessibility resulting from exposure to blue light and also investigates possible changes due to the additional exposure to an external magnetic field. Of special interest are the conformational changes observed in the phosphate-binding loop and the protrusion motif which are both less solvent-exposed upon exposure to blue light.

The results presented in Chapter 5 are already published to a large extent [78]. They introduce the reader to the idea that dimerisation or even higher-order oligomerisation might play a role in *Er*Cry4a magnetoreception exploring non-covalently and covalently linked versions of dimers. The work on this project was interdisciplinary and performed across several research groups. Chapter 5 focuses on the experimental exploration of covalently bound dimers which was carried out by the author. It highlights that cysteine residues C317 and C412 are the most likely candidates for disulphide bonding between monomer units and the fact that dimerisation is promoted upon blue light exposure.

Chapter 6 explores the capabilities of MD simulations as a complementary method to support and fill in the gaps in HDX-MS data. The experimental data were previously obtained by Shane A. Chandler and Katherine Buckton using *Dm*Cry. Therefore, this chapter focuses on the MD simulations of the same protein contributed by the author.

Finally, Chapter 7 summarises the findings and provides an outlook and recommendations for future experiments that were partly already planned during the work on this thesis, but were unfortunately not carried out due to frequent equipment failures.

References

1. Alerstam, T., Hedenström, A. & Åkesson, S. Long-distance migration: evolution and determinants. *Oikos* **103,** 247–260. http://doi.wiley.com/10.1034/j.1600-0706.2003.12559.x (2003).

2. Hore, P. J. & Mouritsen, H. The Radical-Pair Mechanism of Magnetoreception. *Annual Review of Biophysics* **45,** 299–344 (2016).

3. Kirschvink, J. Magnetite-based magnetoreception. *Current Opinion in Neurobiology* **11,** 462–467. https://linkinghub.elsevier.com/retrieve/pii/S095943880000235X (2001).

4. Gagliardo, A. Forty years of olfactory navigation in birds. *Journal of Experimental Biology* **216,** 2165–2171. https://journals.biologists.com/jeb/article/216/12/2165/11389/Forty-years-of-olfactory-navigation-in-birds (2013).

5. Ritz, T., Adem, S. & Schulten, K. A model for photoreceptor-based magnetoreception in birds. *Biophysical Journal* **78,** 707–718. http://dx.doi.org/10.1016/S0006-3495(00)76629-X (2000).

6. Holland, R. A. True navigation in birds: from quantum physics to global migration. *Journal of Zoology* **293,** 1–15. https://onlinelibrary.wiley.com/doi/10.1111/jzo.12107 (2014).

7. Merrill, R. T. & McElhinny, M. W. The Earth's Magnetic Field. Its History, Origin and Planetary Perspective. *J. Fluid Mech.* **156,** 532–536. https://www.cambridge.org/core/product/identifier/S0016756800031630/type/journal_article (1985).

8. Wiltschko, W. & Wiltschko, R. Magnetic orientation and magnetoreception in birds and other animals. *Journal of Comparative Physiology A* **191,** 675–693. http://link.springer.com/10.1007/s00359-005-0627-7 (2005).

9. Mouritsen, H. in *Sturkie's Avian Physiology* 113–133 (Elsevier, 2015). https://linkinghub.elsevier.com/retrieve/pii/B9780124071605000087.

10. Walker, M. M., Dennis, T. E. & Kirschvink, J. L. The magnetic sense and its use in long-distance navigation by animals. *Current Opinion in Neurobiology* **12,** 735–744. https://linkinghub.elsevier.com/retrieve/pii/S0959438802003896 (2002).

11. And others Galizia, C Giovanni and Lledo, P.-M. *Neurosciences-From Molecule to Behavior: a university textbook.* (Springer, 2013).

12. Deguen, R. & Lasbleis, M. in *CISM International Centre for Mechanical Sciences, Courses and Lectures* 2011, 129–212 (Springer, 2020). arXiv: 1905.10416. http://link.springer.com/10.1007/978-3-030-22074-7_5.

13. Herrmann, F. & Vorbach, T. The geodynamo for non-geophysicists. *European Journal of Physics* **41,** 045803. https://iopscience.iop.org/article/10.1088/1361-6404/ab8780 (2020).

14. Wiltschko, R. *Magnetic orientation in animals* (Springer Science & Business Media, 2012).

15. Zadeh-Haghighi, H. & Simon, C. Magnetic field effects in biology from the perspective of the radical pair mechanism. *Journal of The Royal Society Interface* **19.** arXiv: 2204.09147. https://royalsocietypublishing.org/doi/10.1098/rsif.2022.0325 (2022).

16. Alken, P. *et al.* International Geomagnetic Reference Field: the thirteenth generation. *Earth, Planets and Space* **73,** 49. https://doi.org/10.1186/s40623-020-01288-x%20https://earth-planets-space.springeropen.com/articles/10.1186/s40623-020-01288-x (2021).

17. Mouritsen, H. Long-distance navigation and magnetoreception in migratory animals. *Nature* **558,** 50–59. http://www.nature.com/articles/s41586-018-0176-1%20https://www.nature.com/articles/s41586-018-0176-1 (2018).

18. Lohmann, K. J., Putman, N. F. & Lohmann, C. M. The magnetic map of hatchling loggerhead sea turtles. *Current Opinion in Neurobiology* **22,** 336–342. http://dx.doi.org/10.1016/j.conb.2011.11.005%20https://linkinghub.elsevier.com/retrieve/pii/S0959438811001954 (2012).

19. Mouritsen, H. in *Avian Migration* 493–513 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2003). http://link.springer.com/10.1007/978-3-662-05957-9_34.

20. Giorgi, G., Guerra, D., Pezzoli, C., Cavicchi, S. & Bersani, F. Genetic effects of static magnetic fields. Body size increase and lethal mutations induced in populations of Drosophila melanogaster after chronic exposure. *Genetics Selection Evolution* **24,** 393. http://www.gsejournal.org/content/24/5/393 (1992).

21. Ahmad, M., Galland, P., Ritz, T., Wiltschko, R. & Wiltschko, W. Magnetic intensity affects cryptochrome-dependent responses in Arabidopsis thaliana. *Planta* **225,** 615–624 (2007).

22. Harris, S.-R. *et al.* Effect of magnetic fields on cryptochrome-dependent responses in Arabidopsis thaliana. *Journal of The Royal Society Interface* **6,** 1193–1205. https://royalsocietypublishing.org/doi/10.1098/rsif.2008.0519 (2009).

23. Wiltschko, R. & Wiltschko, W. Magnetoreception: Why Is conditioning so seldom successful? *Naturwissenschaften* **83,** 241–247 (1996).

24. Kirschvink, J. L., Winklhofer, M. & Walker, M. M. Biophysics of magnetic orientation: strengthening the interface between theory and experimental design. *Journal of The Royal Society Interface* **7.** https://royalsocietypublishing.org/doi/10.1098/rsif.2009.0491.focus (2010).

25. Bradlaugh, A., Munro, A. L., Jones, A. R. & Baines, R. A. Exploiting the Fruitfly, Drosophila melanogaster, to Identify the Molecular Basis of Cryptochrome-Dependent Magnetosensitivity. *Quantum Reports* **3,** 127–136. https://www.mdpi.com/2624-960X/3/1/7 (2021).

26. Kirschvink, J. L. & Kirschvink, A. K. Is Geomagnetic Sensitivity Real? Replication of the Walker-Bitterman Magnetic Conditioning Experiment in Honey Bees. *American Zoologist* **31,** 169–186. https://academic.oup.com/icb/article-lookup/doi/10.1093/icb/31.1.169 (1991).

27. Lohmann, K. J. & Lohmann, C. A Light-Independent Magnetic Compass in the Leatherback Sea Turtle. *The Biological Bulletin* **185,** 149–151. https://www.journals.uchicago.edu/doi/10.2307/1542138 (1993).

28. Bottesch, M. *et al.* A magnetic compass that might help coral reef fish larvae return to their natal reef. *Current Biology* **26,** R1266–R1267. http://dx.doi.org/10.1016/j.cub.2016.10.051%20https://linkinghub.elsevier.com/retrieve/pii/S0960982216312799 (2016).

29. Brower, L. P. Monarch Butterfly Orientation: Missing Pieces of a Magnificent Puzzle. *Journal of Experimental Biology* **199,** 93–103. https://journals.biologists.com/jeb/article/199/1/93/7381/Monarch-Butterfly-Orientation-Missing-Pieces-of-a (1996).

30. Holland, R. A., Thorup, K., Vonhof, M. J., Cochran, W. W. & Wikelski, M. Bat orientation using Earth's magnetic field. *Nature* **444,** 702–702. https://www.nature.com/articles/444702a (2006).

31. Phillips, J. B. & Borland, S. C. Behavioural evidence for use of a light-dependent magnetoreception mechanism by a vertebrate. *Nature* **359,** 142–144. https://www.nature.com/articles/359142a0 (1992).

32. Foley, L. E., Gegear, R. J. & Reppert, S. M. Human cryptochrome exhibits light-dependent magnetosensitivity. *Nature Communications* **2,** 356. https://www.nature.com/articles/ncomms1364 (2011).

33. Chae, K.-S., Oh, I.-T., Lee, S.-H. & Kim, S.-C. Blue light-dependent human magnetoreception in geomagnetic food orientation. *PLOS ONE* **14** (ed Ritz, T.) e0211826. https://dx.plos.org/10.1371/journal.pone.0211826 (2019).

34. Wiltschko, W. & Wiltschko, R. in *Orientation in birds* 16–37 (1991). https://link.springer.com/10.1007/978-3-0348-7208-9_2.

35. Wiltschko, W. & Wiltschko, R. Migratory Orientation: Magnetic Compass Orientation of Garden Warblers (Sylvia borin) after a Simulated Crossing of the Magnetic Equator. *Ethology* **91,** 70–74 (1992).

36. Berthold, P. A comprehensive theory for the evolution, control and adaptability of avian migration. *Ostrich* **70,** 1–11. https://www.tandfonline.com/doi/full/10.1080/00306525.1999.9639744 (1999).

37. Chernetsov, N. *et al.* Migratory Eurasian Reed Warblers Can Use Magnetic Declination to Solve the Longitude Problem. *Current Biology* **27,** 2647–2651.e2. http://dx.doi.org/10.1016/j.cub.2017.07.024%20https://linkinghub.elsevier.com/retrieve/pii/S0960982217308825 (2017).

38. Wynn, J. *et al.* Magnetic stop signs signal a European songbird's arrival at the breeding site after migration. *Science* **375,** 446–449. https://www.science.org/doi/10.1126/science.abj4210 (2022).

39. Muheim, R., Schmaljohann, H. & Alerstam, T. Feasibility of sun and magnetic compass mechanisms in avian long-distance migration. *Movement Ecology* **6,** 8. https://movementecologyjournal.biomedcentral.com/articles/10.1186/s40462-018-0126-4 (2018).

40. Wiltschko, W. & Wiltschko, R. Magnetic Compass of European Robins. *Science* **176,** 62–64. https://www.science.org/doi/10.1126/science.176.4030.62 (1972).

41. Chernetsov, N., Pakhomov, A., Davydov, A., Cellarius, F. & Mouritsen, H. No evidence for the use of magnetic declination for migratory navigation in two songbird species. *PLOS ONE* **15** (ed Solov'yov, I.) e0232136. http://dx.doi.org/10.1371/journal.pone.0232136%20https://dx.plos.org/10.1371/journal.pone.0232136 (2020).

42. Helm, B. & Muheim, R. Bird migration: Clock and compass facilitate hemisphere switching. *Current Biology* **31,** R1058–R1061. https://doi.org/10.1016/j.cub.2021.07.026%20https://linkinghub.elsevier.com/retrieve/pii/S0960982221009817 (2021).

43. Wiltschko, R., Stapput, K., Thalau, P. & Wiltschko, W. Directional orientation of birds by the magnetic field under different light conditions. *Journal of The Royal Society Interface* **7.** https://royalsocietypublishing.org/doi/10.1098/rsif.2009.0367.focus (2010).

44. Newton, I. *The migration ecology of birds* (Elsevier, 2010).

45. Rogers, L. J., Munro, U., Freire, R., Wiltschko, R. & Wiltschko, W. Lateralized response of chicks to magnetic cues. *Behavioural Brain Research* **186,** 66–71. https://linkinghub.elsevier.com/retrieve/pii/S0166432807003890 (2008).

46. Wiltschko, W., Traudt, J., Güntürkün, O., Wiltschko, R. & Prior, H. Lateralization of magnetic compass orientation in pigeons. *Journal of The Royal Society Interface* **7,** 467–470. https://royalsocietypublishing.org/doi/10.1098/rsif.2009.0436.focus (2010).

47. Hein, C. M., Engels, S., Kishkinev, D. & Mouritsen, H. Robins have a magnetic compass in both eyes. *Nature* **471,** E1–E1. http://dx.doi.org/10.1038/nature09875%20https://www.nature.com/articles/nature09875 (2011).

48. Wiltschko, W., Traudt, J., Güntürkün, O., Prior, H. & Wiltschko, R. Wiltschko et al. reply. *Nature* **471,** E1–E1. https://www.nature.com/articles/nature09876 (2011).

49. Engels, S., Hein, C. M., Lefeldt, N., Prior, H. & Mouritsen, H. Night-Migratory Songbirds Possess a Magnetic Compass in Both Eyes. *PLoS ONE* **7** (ed Bartell, P. A.) e43271. https://dx.plos.org/10.1371/journal.pone.0043271 (2012).

50. Wiltschko, W. & Wiltschko, R. The effect of yellow and blue light on magnetic compass orientation in European robins, Erithacus rubecula. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology* **184,** 295–299. http://link.springer.com/10.1007/s003590050327 (1999).

51. Möller, A., Sagasser, S., Wiltschko, W. & Schierwater, B. Retinal cryptochrome in a migratory passerine bird: a possible transducer for the avian magnetic compass. *Naturwissenschaften* **91,** 585–588. http://link.springer.com/10.1007/s00114-004-0578-9 (2004).

52. Günther, A. *et al.* Double-Cone Localization and Seasonal Expression Pattern Suggest a Role in Magnetoreception for European Robin Cryptochrome 4. *Current Biology* **28,** 211–223.e4 (2018).

53. Einwich, A., Dedek, K., Seth, P. K., Laubinger, S. & Mouritsen, H. A novel isoform of cryptochrome 4 (Cry4b) is expressed in the retina of a night-migratory songbird. *Scientific Reports* **10,** 15794. https://doi.org/10.1038/s41598-020-72579-2%20http://www.nature.com/articles/s41598-020-72579-2%20https://www.nature.com/articles/s41598-020-72579-2 (2020).

54. Einwich, A. *et al.* Localisation of cryptochrome 2 in the avian retina. *Journal of Comparative Physiology A* **208,** 69–81. https://doi.org/10.1007/s00359-021-01506-1%20https://link.springer.com/10.1007/s00359-021-01506-1 (2022).

55. Gagliardo, A. *et al.* Oceanic navigation in Cory's shearwaters: evidence for a crucial role of olfactory cues for homing after displacement. *Journal of Experimental Biology* **216,** 2798–2805. https://journals.biologists.com/jeb/article/216/15/2798/11490/Oceanic-navigation-in-Cory-s-shearwaters-evidence (2013).

56. Murray, R. W. The response of the ampullae of Lorenzini of elasmobranchs to electrical stimulation. *Journal of Experimental Biology* **39,** 119–128 (1962).

57. Paulin, M. G. Electroreception and the compass sense of sharks. *Journal of Theoretical Biology* **174,** 325–339. https://linkinghub.elsevier.com/retrieve/pii/S0022519385701026 (1995).

58. Kalmijn, A. J. Electric and Magnetic Field Detection in Elasmobranch Fishes. *Science* **218,** 916–918. https://www.science.org/doi/10.1126/science.7134985 (1982).

59. Nimpf, S. *et al.* A Putative Mechanism for Magnetoreception by Electromagnetic Induction in the Pigeon Inner Ear. *Current Biology* **29,** 4052–4059.e4. https://doi.org/10.1016/j.cub.2019.09.048%20https://linkinghub.elsevier.com/retrieve/pii/S0960982219312394 (2019).

60. Hochstoeger, T. *et al.* The biophysical, molecular, and anatomical landscape of pigeon CRY4 : A candidate light-based quantal magnetosensor. *Science Advances* **6,** eabb9110 (2020).

61. Wiltschko, R. & Wiltschko, W. Magnetoreception. *BioEssays* **28,** 157–168. https://onlinelibrary.wiley.com/doi/10.1002/bies.20363 (2006).

62. Treiber, C. D. *et al.* High resolution anatomical mapping confirms the absence of a magnetic sense system in the rostral upper beak of pigeons. *Communicative & Integrative Biology* **6,** e24859. http://www.tandfonline.com/doi/abs/10.4161/cib.24859 (2013).

63. Mouritsen, H., Heyers, D. & Güntürkün, O. The Neural Basis of Long-Distance Navigation in Birds. *Annual Review of Physiology* **78,** 133–154. https://www.annualreviews.org/doi/10.1146/annurev-physiol-021115-105054 (2016).

64. Binhi, V. N. Nonspecific magnetic biological effects: A model assuming the spin-orbit coupling. *The Journal of Chemical Physics* **151.** https://doi.org/10.1063/1.5127972%20https://pubs.aip.org/jcp/article/151/20/204101/197270/Nonspecific-magnetic-biological-effects-A-model (2019).

65. Dhiman, S. K., Wu, F. & Galland, P. Effects of weak static magnetic fields on the development of seedlings of Arabidopsis thaliana. *Protoplasma* **260,** 767–786. https://doi.org/10.1007/s00709-022-01811-9%20https://link.springer.com/10.1007/s00709-022-01811-9 (2023).

66. Woodward, J. R. Radical Pairs in Solution. *Progress in Reaction Kinetics and Mechanism* **27,** 165–207. http://journals.sagepub.com/doi/10.3184/007967402103165388 (2002).

67. Schulten, K., Swenberg, C. E. & Weller, A. A biomagnetic sensory mechanism based on magnetic field modulated coherent electron spin motion. *Zeitschrift für Physikalische Chemie* **111,** 1–5. https://www.degruyter.com/document/doi/10.1524/zpch.1978.111.1.001/html (1978).

68. Liedvogel, M. & Mouritsen, H. Cryptochromes—a potential magnetoreceptor: what do we know and what do we want to know? *Journal of The Royal Society Interface* **7.** https://royalsocietypublishing.org/doi/10.1098/rsif.2009.0411.focus (2010).

69. Liu, Z. *et al.* Determining complete electron flow in the cofactor photoreduction of oxidized photolyase. *Proceedings of the National Academy of Sciences* **110,** 12966–12971. https://pnas.org/doi/full/10.1073/pnas.1311073110 (2013).

70. Solov'Yov, I. A., Domratcheva, T., Moughal Shahi, A. R. & Schulten, K. Decrypting cryptochrome: Revealing the molecular identity of the photoactivation reaction. *Journal of the American Chemical Society* **134,** 18046–18052 (2012).

71. Xu, J. *et al.* Magnetic sensitivity of cryptochrome 4 from a migratory songbird. *Nature* **594,** 535–540. http://dx.doi.org/10.1038/s41586-021-03618-9 (2021).

72. Zoltowski, B. D. *et al.* Chemical and structural analysis of a photoactive vertebrate cryptochrome from pigeon. *Proceedings of the National Academy of Sciences* **116,** 19449–19457. http://www.pnas.org/lookup/doi/10.1073/pnas.1907875116 (2019).

73. Franz, S. *et al.* Structure of the bifunctional cryptochrome aCRY from Chlamydomonas reinhardtii. *Nucleic Acids Research* **46,** 8010–8022. https://academic.oup.com/nar/article/46/15/8010/5055825 (2018).

74. Nohr, D. *et al.* Extended Electron-Transfer in Animal Cryptochromes Mediated by a Tetrad of Aromatic Amino Acids. *Biophysical Journal* **111,** 301–311. http://dx.doi.org/10.1016/j.bpj.2016.06.009%20https://linkinghub.elsevier.com/retrieve/pii/S0006349516304428 (2016).

75. Hong, G., Pachter, R., Essen, L. O. & Ritz, T. Electron transfer and spin dynamics of the radical-pair in the cryptochrome from Chlamydomonas reinhardtii by computational analysis. *Journal of Chemical Physics* **152.** https://doi.org/10.1063/1.5133019 (2020).

76. Chaves, I. *et al.* The Cryptochromes: Blue Light Photoreceptors in Plants and Animals. *Annual Review of Plant Biology* **62,** 335–364. https://www.annualreviews.org/doi/10.1146/annurev-arplant-042110-103759 (2011).

77. Michael, A. K., Fribourgh, J. L., Van Gelder, R. N. & Partch, C. L. Animal Cryptochromes: Divergent Roles in Light Perception, Circadian Timekeeping and Beyond. *Photochemistry and Photobiology* **93,** 128–140. https://onlinelibrary.wiley.com/doi/10.1111/php.12677 (2017).

78. Hanić, M. *et al.* Dimerization of European Robin Cryptochrome 4a. *The Journal of Physical Chemistry B* **127,** 6251–6264. https://pubs.acs.org/doi/10.1021/acs.jpcb.3c01305 (2023).

79. Ritz, T., Thalau, P., Phillips, J. B., Wiltschko, R. & Wiltschko, W. Resonance effects indicate a radical-pair mechanism for avian magnetic compass. *Nature* **429,** 177–180. https://www.nature.com/articles/nature02534 (2004).

80. Engels, S. *et al.* Anthropogenic electromagnetic noise disrupts magnetic compass orientation in a migratory bird. *Nature* **509,** 353–356. https://www.nature.com/articles/nature13290 (2014).

81. Hiscock, H. G., Mouritsen, H., Manolopoulos, D. E. & Hore, P. Disruption of Magnetic Compass Orientation in Migratory Birds by Radiofrequency Electromagnetic Fields. *Biophysical Journal* **113,** 1475–1484. https://doi.org/10.1016/j.bpj.2017.07.031%20https://linkinghub.elsevier.com/retrieve/pii/S0006349517308597 (2017).

82. Leberecht, B. *et al.* Broadband 75–85 MHz radiofrequency fields disrupt magnetic compass orientation in night-migratory songbirds consistent with a flavin-based radical pair magnetoreceptor. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology* **208,** 97–106. https://doi.org/10.1007/s00359-021-01537-8 (2022).

83. Leberecht, B. *et al.* Upper bound for broadband radiofrequency field disruption of magnetic compass orientation in night-migratory songbirds. *Proceedings of the National Academy of Sciences* **120,** 2017. http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2216830120/-/DCSupplemental.https://doi.org/10.1073/pnas.2216830120%20https://pnas.org/doi/10.1073/pnas.2301153120 (2023).

84. Ritz, T. *et al.* Magnetic Compass of Birds Is Based on a Molecule with Optimal Directional Sensitivity. *Biophysical Journal* **96,** 3451–3457. http://dx.doi.org/10.1016/j.bpj.2008.11.072%20https://linkinghub.elsevier.com/retrieve/pii/S0006349509004688 (2009).

85. Schwarze, S. *et al.* Weak Broadband Electromagnetic Fields are More Disruptive to Magnetic Compass Orientation in a Night-Migratory Songbird (Erithacus rubecula) than Strong Narrow-Band Fields. *Frontiers in Behavioral Neuroscience* **10,** 1–13. http://journal.frontiersin.org/Article/10.3389/fnbeh.2016.00055/abstract (2016).

86. Lee, A. A. *et al.* Alternative radical pairs for cryptochrome-based magnetoreception. *Journal of The Royal Society Interface* **11,** 20131063. https://royalsocietypublishing.org/doi/10.1098/rsif.2013.1063 (2014).

87. Babcock, N. S. & Kattnig, D. R. Radical Scavenging Could Answer the Challenge Posed by Electron–Electron Dipolar Interactions in the Cryptochrome Compass Model. *JACS Au* **1,** 2033–2046. https://pubs.acs.org/doi/10.1021/jacsau.1c00332 (2021).

88. Brudler, R. *et al.* Identification of a New Cryptochrome Class. *Molecular Cell* **11,** 59–67. https://linkinghub.elsevier.com/retrieve/pii/S109727650300008X (2003).

89. Navarro, E. *et al.* The DASH-type Cryptochrome from the Fungus Mucor circinelloides Is a Canonical CPD-Photolyase. *Current Biology* **30,** 4483–4490.e4. https://linkinghub.elsevier.com/retrieve/pii/S0960982220312410 (2020).

90. Ozturk, N. Phylogenetic and Functional Classification of the Photolyase/Cryptochrome Family. *Photochemistry and Photobiology* **93,** 104–111 (2017).

91. Kavet, R. & Brain, J. Cryptochromes in Mammals and Birds: Clock or Magnetic Compass? *Physiology* **36,** 183–194. https://journals.physiology.org/doi/10.1152/physiol.00040.2020 (2021).

92. Parico, G. C. G. & Partch, C. L. The tail of cryptochromes: an intrinsically disordered cog within the mammalian circadian clock. *Cell Communication and Signaling* **18,** 182. https://biosignaling.biomedcentral.com/articles/10.1186/s12964-020-00665-z (2020).

93. Sancar, A. Structure and Function of DNA Photolyase and Cryptochrome Blue-Light Photoreceptors. *Chemical Reviews* **103,** 2203–2238. https://pubs.acs.org/doi/10.1021/cr0204348 (2003).

94. Berndt, A. *et al.* A Novel Photoreaction Mechanism for the Circadian Blue Light Photoreceptor Drosophila Cryptochrome. *Journal of Biological Chemistry* **282,** 13011–13021. http://dx.doi.org/10.1074/jbc.M608872200%20https://linkinghub.elsevier.com/retrieve/pii/S0021925818950751 (2007).

95. Lin, C. & Todo, T. The cryptochromes. *Genome Biology* **6,** 1–9 (2005).

96. Ozturk, N. Light-dependent reactions of animal circadian photoreceptor cryptochrome. *The FEBS Journal* **289,** 6622–6639. https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16273 (2022).

97. Gegear, R. J., Foley, L. E., Casselman, A. & Reppert, S. M. Animal cryptochromes mediate magnetoreception by an unconventional photochemical mechanism. *Nature* **463,** 804–807. https://www.nature.com/articles/nature08719 (2010).

98. Gressel, J. Yearly review: Blue light photoreception. *Photochemistry and Photobiology* **30,** 749–754 (1979).

99. Malhotra, K., Kim, S.-T., Batschauer, A., Dawut, L. & Sancar, A. Putative Blue-Light Photoreceptors from Arabidopsis thaliana and Sinapis alba with a High Degree of Sequence Homology to DNA Photolyase Contain the Two Photolyase Cofactors but Lack DNA Repair Activity. *Biochemistry* **34,** 6892–6899. https://pubs.acs.org/doi/abs/10.1021/bi00020a037 (1995).

100. Ahmad, M. & Cashmore, A. R. HY4 gene of A. thaliana encodes a protein with characteristics of a blue-light photoreceptor. *Nature* **366,** 162–166. https://www.nature.com/articles/366162a0 (1993).

101. Yu, X., Liu, H., Klejnot, J. & Lin, C. The Cryptochrome Blue Light Receptors. *The Arabidopsis Book* **8,** e0135. arXiv: NIHMS150003. http://www.bioone.org/doi/abs/10.1199/tab.0135 (2010).

102. Lin, C., Ahmad, M., Gordon, D. & Cashmore, A. R. Expression of an Arabidopsis cryptochrome gene in transgenic tobacco results in hypersensitivity to blue, UV-A, and green light. *Proceedings of the National Academy of Sciences* **92,** 8423–8427. https://pnas.org/doi/full/10.1073/pnas.92.18.8423 (1995).

103. Guo, H., Yang, H., Mockler, T. C. & Lin, C. Regulation of Flowering Time by Arabidopsis Photoreceptors. *Science* **279,** 1360–1363. https://journals.biologists.com/dev/article/133/9/1693/52948/Regulation-of-flowering-time-by-Arabidopsis-MSI1%20https://www.science.org/doi/10.1126/science.279.5355.1360 (1998).

104. Kleine, T., Lockhart, P. & Batschauer, A. An Arabidopsis protein closely related to Synechocystis cryptochrome is targeted to organelles. *The Plant Journal* **35,** 93–103. http://doi.wiley.com/10.1046/j.1365-313X.2003.01787.x (2003).

105. Zhang, L. & Malkemper, E. P. Cryptochromes in mammals: a magnetoreception misconception? *Frontiers in Physiology* **14,** 1–9. https://www.frontiersin.org/articles/10.3389/fphys.2023.1250798/full (2023).

106. Nitta, Y. *et al.* Analysing the evolutional and functional differentiation of four types of Daphnia magna cryptochrome in Drosophila circadian clock. *Scientific Reports* **9,** 8857. http://dx.doi.org/10.1038/s41598-019-45410-w%20https://www.nature.com/articles/s41598-019-45410-w (2019).

107. Balay, S. D. *et al.* The expression, localisation and interactome of pigeon CRY2. *Scientific Reports* **11,** 20293. https://doi.org/10.1038/s41598-021-99207-x%20https://www.nature.com/articles/s41598-021-99207-x (2021).

108. Nießner, C. *et al.* Seasonally Changing Cryptochrome 1b Expression in the Retinal Ganglion Cells of a Migrating Passerine Bird. *PLOS ONE* **11** (ed Warrant, E. J.) e0150377. https://dx.plos.org/10.1371/journal.pone.0150377 (2016).

109. Cashmore, A. R., Jarillo, J. A., Wu, Y.-J. & Liu, D. Cryptochromes: Blue Light Receptors for Plants and Animals. *Science* **284,** 760–765. https://www.science.org/doi/10.1126/science.284.5415.760 (1999).

110. Yoshii, T., Ahmad, M. & Helfrich-Förster, C. Cryptochrome Mediates Light-Dependent Magnetosensitivity of Drosophila's Circadian Clock. *PLoS Biology* **7** (ed Schibler, U.) e1000086. https://dx.plos.org/10.1371/journal.pbio.1000086 (2009).

111. Hsu, D. S. *et al.* Putative Human Blue-Light Photoreceptors hCRY1 and hCRY2 Are Flavoproteins. *Biochemistry* **35,** 13871–13877. https://pubs.acs.org/doi/10.1021/bi962209o (1996).

112. Griffin, E. A., Staknis, D. & Weitz, C. J. Light-Independent Role of CRY1 and CRY2 in the Mammalian Circadian Clock. *Science* **286,** 768–771. https://www.science.org/doi/10.1126/science.286.5440.768 (1999).

113. Bradlaugh, A. A. *et al.* Essential elements of radical pair magnetosensitivity in Drosophila. *Nature* **615,** 111–116. https://www.biorxiv.org/content/10.1101/2021.10.29.466426v1%0Ahttps://www.biorxiv.org/content/10.1101/2021.10.29.466426v1.abstract%20https://www.nature.com/articles/s41586-023-05735-z (2023).

114. Fogle, K. J. *et al.* CRYPTOCHROME-mediated phototransduction by modulation of the potassium ion channel β-subunit redox sensor. *Proceedings of the National Academy of Sciences* **112,** 2245–2250. https://pnas.org/doi/full/10.1073/pnas.1416586112 (2015).

115. Wu, H., Scholten, A., Einwich, A., Mouritsen, H. & Koch, K.-W. Protein-protein interaction of the putative magnetoreceptor cryptochrome 4 expressed in the avian retina. *Scientific Reports* **10,** 7364. http://www.nature.com/articles/s41598-020-64429-y%20https://www.nature.com/articles/s41598-020-64429-y (2020).

116. Görtemaker, K. *et al.* Direct Interaction of Avian Cryptochrome 4 with a Cone Specific G-Protein. *Cells* **11,** 2043. https://www.mdpi.com/2073-4409/11/13/2043 (2022).

117. Qin, S. *et al.* A magnetic protein biocompass. *Nature Materials* **15,** 217–226. https://www.nature.com/articles/nmat4484 (2016).

118. Partch, C. L., Clarkson, M. W., Özgür, S., Lee, A. L. & Sancar, A. Role of Structural Plasticity in Signal Transduction by the Cryptochrome Blue-Light Photoreceptor. *Biochemistry* **44,** 3795–3805. https://pubs.acs.org/doi/10.1021/bi047545g (2005).

119. Wang, H., Ma, L.-G., Li, J.-M., Zhao, H.-Y. & Deng, X. W. Direct Interaction of Arabidopsis Cryptochromes with COP1 in Light Control Development. *Science* **294,** 154–158. https://www.science.org/doi/10.1126/science.1063630 (2001).

120. Yang, H.-Q. The Signaling Mechanism of Arabidopsis CRY1 Involves Direct Interaction with COP1. *The Plant Cell* **13,** 2573–2587. http://www.plantcell.org/cgi/doi/10.1105/tpc.13.12.2573 (2001).

121. Kondoh, M. *et al.* Light-Induced Conformational Changes in Full-Length Arabidopsis thaliana Cryptochrome. *Journal of Molecular Biology* **413,** 128–137. http://dx.doi.org/10.1016/j.jmb.2011.08.031%20https://linkinghub.elsevier.com/retrieve/pii/S0022283611009296 (2011).

122. Sang, Y. *et al.* N-Terminal Domain–Mediated Homodimerization Is Required for Photoreceptor Activity of Arabidopsis CRYPTOCHROME 1. *The Plant Cell* **17,** 1569–1584. https://academic.oup.com/plcell/article/17/5/1569-1584/6114536 (2005).

123. Park, H. *et al.* Optogenetic protein clustering through fluorescent protein tagging and extension of CRY2. *Nature Communications* **8,** 30. http://dx.doi.org/10.1038/s41467-017-00060-2%20http://www.nature.com/articles/s41467-017-00060-2%20https://www.nature.com/articles/s41467-017-00060-2 (2017).

124. Watari, R. *et al.* Light-dependent Structural Change of Chicken Retinal Cryptochrome4. *Journal of Biological Chemistry* **287,** 42634–42641. https://linkinghub.elsevier.com/retrieve/pii/S0021925820437561 (2012).

125. Krischer, J., König, S., Weisheit, W., Mittag, M. & Büchel, C. The C-Terminus of a diatom plant-like cryptochrome influences the FAD redox state and binding of interaction partners. *Journal of Experimental Botany* **73,** 1934–1948 (2022).

126. Zoltowski, B. D. *et al.* Structure of full-length Drosophila cryptochrome. *Nature* **480,** 396–399. https://www.nature.com/articles/nature10618 (2011).

127. Levy, C. *et al.* Updated structure of Drosophila cryptochrome. *Nature* **495,** E3–E4. http://www.nature.com/articles/nature11995 (2013).

128. Czarna, A. *et al.* Structures of Drosophila Cryptochrome and Mouse Cryptochrome1 Provide Insight into Circadian Function. *Cell* **153,** 1394–1405. http://dx.doi.org/10.1016/j.cell.2013.05.011%20https://linkinghub.elsevier.com/retrieve/pii/S009286741300576X (2013).

129. Vaidya, A. T. *et al.* Flavin reduction activates Drosophila cryptochrome. *Proceedings of the National Academy of Sciences* **110,** 20455–20460. https://pnas.org/doi/full/10.1073/pnas.1313336110 (2013).

130. Chandrasekaran, S. *et al.* Tuning flavin environment to detect and control light-induced conformational switching in Drosophila cryptochrome. *Communications Biology* **4,** 249. http://dx.doi.org/10.1038/s42003-021-01766-2%20http://www.nature.com/articles/s42003-021-01766-2%20https://www.nature.com/articles/s42003-021-01766-2 (2021).

131. Fedele, G., Green, E. W., Rosato, E. & Kyriacou, C. P. An electromagnetic field disrupts negative geotaxis in Drosophila via a CRY-dependent pathway. *Nature Communications* **5,** 4391. http://www.nature.com/articles/ncomms5391 (2014).

132. Gegear, R. J., Casselman, A., Waddell, S. & Reppert, S. M. Cryptochrome mediates light-dependent magnetosensitivity in Drosophila. *Nature* **454,** 1014–1018 (2008).

133. Bassetto, M. *et al.* No evidence for magnetic field effects on the behaviour of Drosophila. *Nature* **620,** 595–599. https://www.nature.com/articles/s41586-023-06397-7 (2023).

134. Fedele, G. *et al.* Genetic Analysis of Circadian Responses to Low Frequency Electromagnetic Fields in Drosophila melanogaster. *PLoS Genetics* **10** (ed Taghert, P. H.) e1004804. https://dx.plos.org/10.1371/journal.pgen.1004804 (2014).

135. Hemsley, M. J. *et al.* Linear motifs in the C-terminus of D. melanogaster cryptochrome. *Biochemical and Biophysical Research Communications* **355,** 531–537. https://linkinghub.elsevier.com/retrieve/pii/S0006291X07002690 (2007).

136. Mazzotta, G. M. *et al.* Calmodulin Enhances Cryptochrome Binding to INAD in Drosophila Photoreceptors. *Frontiers in Molecular Neuroscience* **11,** 1–15. https://www.frontiersin.org/article/10.3389/fnmol.2018.00280/full (2018).

137. Kattnig, D. R., Nielsen, C. & Solov'yov, I. A. Molecular dynamics simulations disclose early stages of the photo-activation of cryptochrome 4. *New Journal of Physics* **20,** 083018. https://iopscience.iop.org/article/10.1088/1367-2630/aad70f (2018).

138. Schuhmann, F., Kattnig, D. R. & Solov'yov, I. A. Exploring Post-activation Conformational Changes in Pigeon Cryptochrome 4. *Journal of Physical Chemistry B* **125,** 9652–9659 (2021).

139. *Protein dimerization and oligomerization in biology* (ed Matthews, J. M.) (Springer Science & Business Media, 2012).

140. Heldin, C.-H. Dimerization of cell surface receptors in signal transduction. *Cell* **80,** 213–223. https://linkinghub.elsevier.com/retrieve/pii/0092867495904042 (1995).

141. Palayam, M. *et al.* Structural insights into photoactivation of plant Cryptochrome-2. *Communications Biology* **4,** 28. http://dx.doi.org/10.1038/s42003-020-01531-x%20https://www.nature.com/articles/s42003-020-01531-x (2021).

142. Shao, K. *et al.* The oligomeric structures of plant cryptochromes. *Nature Structural & Molecular Biology* **27,** 480–488. http://www.nature.com/articles/s41594-020-0420-x (2020).

143. Poehn, B. *et al.* A Cryptochrome adopts distinct moon- and sunlight states and functions as sun- versus moonlight interpreter in monthly oscillator entrainment. *Nature Communications* **13,** 5220. https://www.nature.com/articles/s41467-022-32562-z (2022).

144. Oldemeyer, S. *et al.* Essential Role of an Unusually Long-lived Tyrosyl Radical in the Response to Red Light of the Animal-like Cryptochrome aCRY. *Journal of Biological Chemistry* **291,** 14062–14071. http://dx.doi.org/10.1074/jbc.M116.726976%20https://linkinghub.elsevier.com/retrieve/pii/S0021925820367545 (2016).

145. Lindner, R., Heintz, U. & Winkler, A. Applications of hydrogen deuterium exchange (HDX) for the characterization of conformational dynamics in light-activated photoreceptors. *Frontiers in Molecular Biosciences* **2,** 1–13. http://journal.frontiersin.org/Article/10.3389/fmolb.2015.00033/abstract (2015).

146. Hoff, W. D. *et al.* Global Conformational Changes upon Receptor Stimulation in Photoactive Yellow Protein. *Biochemistry* **38,** 1009–1017. https://pubs.acs.org/doi/10.1021/bi980504y (1999).

147. Brudler, R. *et al.* PAS Domain Allostery and Light-induced Conformational Changes in Photoactive Yellow Protein upon I2 Intermediate Formation, Probed with Enhanced Hydrogen/Deuterium Exchange Mass Spectrometry. *Journal of Molecular Biology* **363,** 148–160. https://linkinghub.elsevier.com/retrieve/pii/S0022283606009752 (2006).

148. Craven, C. J. *et al.* Probing the Nature of the Blue-Shifted Intermediate of Photoactive Yellow Protein in Solution by NMR: Hydrogen-Deuterium Exchange Data and pH Studies. *Biochemistry* **39,** 14392–14399. https://pubs.acs.org/doi/10.1021/bi001628p (2000).

149. Von Horsten, S. *et al.* Mapping light-driven conformational changes within the photosensory module of plant phytochrome B. *Scientific Reports* **6,** 34366. https://www.nature.com/articles/srep34366 (2016).

150. Winkler, A. *et al.* A ternary AppA–PpsR–DNA complex mediates light regulation of photosynthesis-related gene expression. *Nature Structural & Molecular Biology* **20,** 859–867. https://www.nature.com/articles/nsmb.2597 (2013).

151. Martens, C. & Politis, A. A glimpse into the molecular mechanism of integral membrane proteins through hydrogen–deuterium exchange mass spectrometry. *Protein Science* **29,** 1285–1301. https://onlinelibrary.wiley.com/doi/10.1002/pro.3853 (2020).

152. Bolla, J. R., Fiorentino, F. & Robinson, C. V. Mass spectrometry informs the structure and dynamics of membrane proteins involved in lipid and drug transport. *Current Opinion in Structural Biology* **70,** 53–60. https://doi.org/10.1016/j.sbi.2021.03.014%20https://linkinghub.elsevier.com/retrieve/pii/S0959440X2100052X (2021).

153. Pan, Y., Brown, L. & Konermann, L. Hydrogen Exchange Mass Spectrometry of Bacteriorhodopsin Reveals Light-Induced Changes in the Structural Dynamics of a Biomolecular Machine. *Journal of the American Chemical Society* **133,** 20237–20244. https://pubs.acs.org/doi/10.1021/ja206197h (2011).

154. Camacho, I. S. *et al.* Native mass spectrometry reveals the conformational diversity of the UVR8 photoreceptor. *Proceedings of the National Academy of Sciences* **116,** 1116–1125. https://pnas.org/doi/full/10.1073/pnas.1813254116 (2019).

155. Camacho, I. S. *et al.* Interplay between chromophore binding and domain assembly by the B 12 -dependent photoreceptor protein, CarH. *Chemical Science* **12,** 8333–8341. http://xlink.rsc.org/?DOI=D1SC00522G (2021).

156. Chen, S. *et al.* Capturing a rhodopsin receptor signalling cascade across a native membrane. *Nature* **604,** 384–390 (2022).

157. Wiltschko, W. & Merkel, F. W. Orientierung zugunruhiger Rotkehlchen im statischen Magnetfeld. *Verh. Dtsch. Zool. Ges* **59,** 362–367.

Most of all, the world is a place where parts of wholes are described within an overarching paradigm of clarity and accuracy. The context in which makes possible an underlying sense of the way it all fits together, despite our collective tendency not to conceive of it as such.

...

So instead, we went ahead to fabricate a catalog of unstable elements and modicums and particles. With not zero total strangeness for brief moments which amount to nothing more than tiny fragments of a finger snap.

— from The Books' *Smells Like Content*

2 Technical Background

Contents

2.1 Mass spectrometry and mass photometry 43
2.1.1 Mass spectrometry instrumentation 44
2.1.2 Solution hydrogen/deuterium exchange mass spectrometry 55
2.1.3 Mass photometry 62
2.2 Electron Paramagnetic Resonance Spectroscopy 64
2.2.1 Continuous wave EPR 65
2.2.2 Pulse EPR 66
2.3 Molecular Dynamics Simulations 69
2.3.1 Force fields and water models 71
2.3.2 Constraints and conditions: periodic boundary conditions, cutoffs, thermo- and barostats 73
References 74

2.1 Mass spectrometry and mass photometry

Mass spectrometry (MS) has many useful applications in biophysics and structural biology. It is a powerful tool used as one of several analytical techniques which often complement each other [1], such as X-ray crystallography, cryo-electron microscopy, small-angle X-ray scattering, nuclear magnetic resonance spectroscopy,

electron paramagnetic resonance spectroscopy, mass photometry (MP), sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE), and sometimes molecular dynamics (MD) simulations.

Here, mostly MS techniques were used to investigate the state of proteins and peptides under different conditions making use of the techniques native mass spectrometry, chemical cross-linking followed by mass spectrometric analysis (XL-MS), and hydrogen/deuterium exchange mass spectrometry (HDX-MS). This chapter therefore mainly focuses on the technical background of MS. Further techniques used in this work and therefore also discussed in this chapter are MP, double electron electron resonance electron paramagnetic resonance (DEER-EPR), and MD simulations.

2.1.1 Mass spectrometry instrumentation

All mass spectrometers combine a few essential elements, an ion source, a vacuum system, a mass analyser and a detector. These basic elements had already been put together in 1918 by A. J. Dempster who wanted to analyse isotopic distributions of metals [2]. Over time, more and more components were added and modified. There are at least six different ionisation sources, different types of mass analysers with varying degrees of resolution and mass accuracy, and similarly different types of detectors. Nowadays, mass analysers are often used in tandem within one instrument to select ions of certain m/z values for subsequent manipulation and analysis.

Mass spectrometry has a wealth of applications and gained more and more popularity over the last decades. Different methods can be employed such as native MS, HDX-MS and the combined use of mass spectrometry and separation techniques such as liquid chromatography MS (LC-MS) or ion mobility MS [3]. Native MS and HDX-MS were the methods of choice for the work presented here. This chapter will therefore initially focus on native MS before moving on to HDX-MS.

Native MS is a powerful technique to elucidate structure and function of large biomolecules in their folded, meaning in their native or near-native, state. Most proteins studied using native MS have been shown to retain their structural integrity

in the gas phase although the extent to which this is possible is still subject to debate [4]. However, native MS has the advantage of being able to investigate dynamic complexes which other, popular structural biology tools struggle to do, such as X-ray crystallography or cryo-electron microscopy. Therefore, native MS has been successfully used as a complementary tool for structural biology approaches, for example to complement interactomics data with information on exact protein complex compositions [5]. Native MS has also gained popularity in the field of drug discovery for screening compounds against target molecules [6, 7]. Further methodological extensions that could be used in combination with native MS to help with this endeavour are tandem MS which is briefly discussed below and ion mobility MS where ions are, in addition to the usual separation by m/z values, also separated by their cross-section dependent drift time [8]. The latter technique was not used for the work presented in this thesis and is therefore not further discussed.

HDX-MS has the ability to probe conformational changes or interaction surfaces of biomolecules by exposing the biomolecules to a label that can only access the currently surface exposed areas at the time of labelling. Subsequent mass analysis can therefore yield information on changes in surface exposed areas. Further information on the capabilities of this technique can be found in section 2.1.2 below.

Electrospray ionisation for native MS

Electrospray ionisation (ESI) is a technique which allows gentle transfer of intact biological molecules into the gas phase for subsequent MS analysis. It is particularly useful for the transferral of larger units such as proteins in their native or near-native states.

By 1984 advances had been made to develop a new kind of ion source using gaseous dispersion of an electrosprayed solution and it was suggested that this method could be applied to larger organic molecules [9, 10]. In 1988, biomolecules within a mass range of 5,000 and 40,000 Da were successfully transferred within a MS setup [11] and by 1990 ESI had been developed for large biological compounds with masses of more than 100,000 Da [12]. For example, Fenn *et al.* were able to

show the successful transmission of biopolymers including proteins with masses of up to 130,000 Da using ESI [13]. Nowadays, large protein complexes consisting of multiple units can be transmitted and analysed with the help of ESI, even those situated in membranes [14–16].

After ESI had been successfully developed, the method was further improved by the development of nano ESI (nESI) which required smaller sample quantities as the orifices of the capillaries were of a smaller diameter. Now only flow rates of a few nL min^{-1} and sample volumes of a few µL were needed [17–19]. A smaller orifice leads to the formation of smaller droplets which increases the sensitivity [20] and lowers the salt:protein ratio which in turn increases the salt tolerance [21, 22].

nESI is a method often employed when conducting native MS measurements where the aim is to transfer the biomolecules through the mass spectrometer without disrupting any of their internal interactions, that is in their native or near-native states, to analyse their structure and function [23]. For this a solution is necessary that facilitates electrospraying of the biological compounds. Ammonium acetate and ammonium bicarbonate are buffers most commonly used for this purpose. Gel filtration or centrifugal membrane columns as well as dialysis are usually used in order to remove salts and transfer the samples into the spraying buffer [24]. The buffers have the advantages that they easily evaporate during ion desolvation, reduce the aggregation of the biomolecules and maintain their structural integrity by ensuring that electrochemical effects within the capillary do not affect the solution pH up to high ionic strengths. The capillary itself usually consists of borosilicate glass, has been pulled to form micro-pipettes and then coated with a conducting material such as gold. The ends of the capillary are cut manually to generate an orifice of 1-5 µm diameter, depending on the instrument and analyte requirements [25]. When preparing a new protein for ESI it should also be taken into account that a folded protein has a net charge which could lead to interactions with the emitter surface, e.g. a positively charged ion could adhere to a negatively charged glass surface of the capillary. The smaller the capillary, the higher the surface to volume ratio and the shorter the diffusion length which increases the risk of adhesion effects [26].

A potential difference between the conducting capillary and the inlet of the mass spectrometer then ensures the transmission of the charged biomolecules into the gas phase [27]. For this a voltage of 1-2 kV is usually applied to the capillary which is positioned a few millimetres away from the source of the mass spectrometer [28]. Figure 2.1A illustrates the following processes: once the charged analyte solution reaches the orifice of the capillary, charged droplets are formed. Several droplet fission events and evaporation steps follow before the charged biomolecules enter the mass spectrometer. During each sequential fission event a jet of offspring droplets emerges [14]. Evaporation of the droplets occurs until the surface tension and the Coulombic repulsion of the charges are equal and the Rayleigh limit is reached [29]. At this stage, fission occurs. The exact mechanism for this is unknown, but it has been proposed to either follow the 'ion evaporation model' where the charged analyte ions are expelled from the droplet [30] or the 'charged residue model' where evaporation and fission occur until only the charged analyte is left [31]. Figure 2.1B illustrates both of these processes. Over time, experimental and theoretical evidence has accumulated that supports the latter theory for globular molecules [32]. A recent publication by Pimlott *et al.* highlights, that there is a third model, the 'chain ejection model', in which a protein gets gradually expelled from a droplet, which has been shown to be the case for unfolded proteins [33].

For the measurement to be successful and constant over several minutes, a stable spray has to be established by forming a Taylor cone from which a constant jet of small droplets is formed [34]. This process can be aided by applying a sheath gas at the capillary which improves the efficiency of the desolvation process [35]. If necessary, a backing pressure of 1-2 bar can be applied at the rear of the capillary to initiate the flow of the sample in solution towards the capillary tip which can subsequently be reduced or turned off once a stable spray has been established [28]. Due to capillary and electrochemical effects it might take a few minutes until the spray leads to a stable signal [36, 37].

The subsequently measured mass spectrum can yield a wealth of information about the system of investigation. Aside from the measured mass over charge

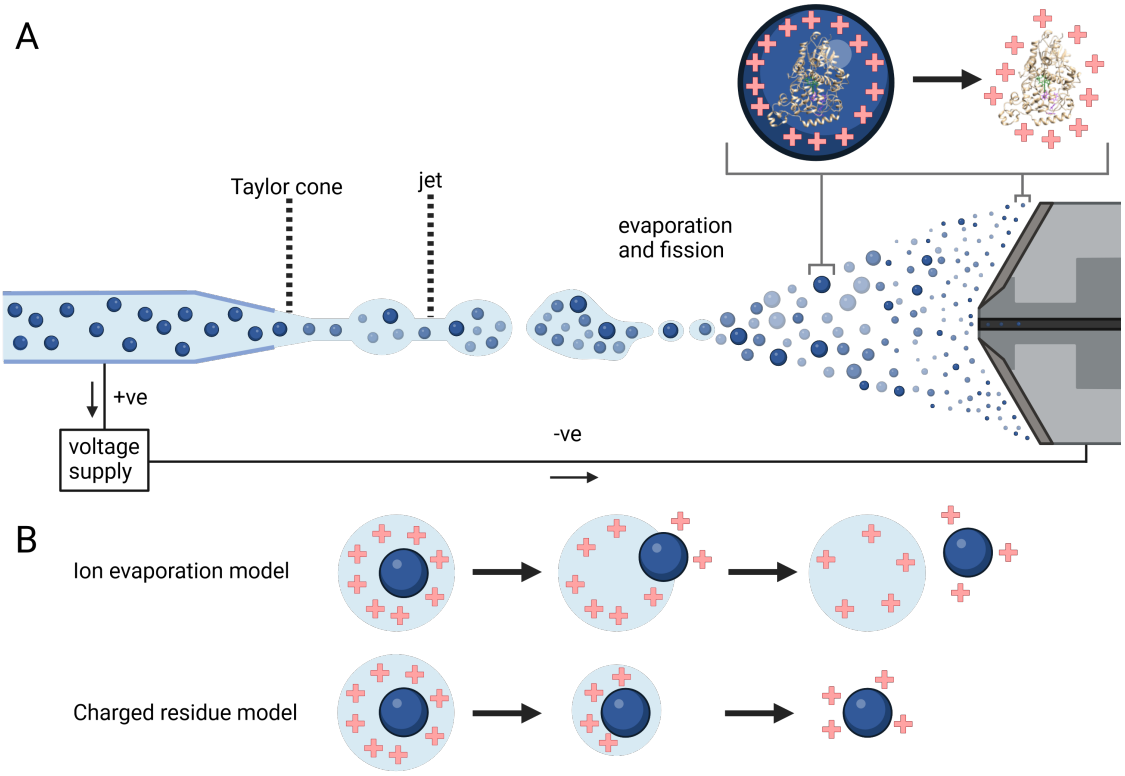

Figure 2.1: Electrospray ionisation of proteins. A A protein solution is sprayed through the orifice of a glass capillary. A high voltage supply is attached to the conducting capillary which charges the solution. A Taylor cone forms at its tip which leads into a thin jet from which droplets form that subsequently undergo several steps of solvent evaporation and fission which occur when the surface charge destabilises the droplet at the Rayleigh limit. This leads to fission and the emergence of small droplets from the initial droplet. **B** Two proposed mechanisms by which the evaporation and fission steps occur. The ion evaporation model assumes that the charged residue will eventually be expelled from the droplet whereas the charged residue model assumes evaporation of all solvent leaving only a charged residue behind. Figure created with BioRender.com.

ratios (m/z) which inform about the different components present in an analyte solution, there is also a correlation between surface area and the average charge state. Therefore, the surface area of a biomolecule can be estimated even without the existence of a crystal structure [38]. Further to this, the width of a charge state distribution can yield information on the integrity of the structure of a biomolecule since less tightly packed structures will allow the accommodation of more charges than would more tightly packed ones [39]. It is also possible to follow the folding and dynamics of a biomolecule as well as to analyse its behaviour

towards interaction partners, e.g. protein-protein interactions, (self-)assemblies or protein-ligand interactions [40].

Transmission of high molecular weight ions

Quadrupole mass analysers, important parts of conventional MS setups which are discussed in further detail below, act as a mass channel for ions before they reach a detector. However, commercial devices usually only cover a m/z range of up to 3000-4000 m/z [41]. Since biomolecules and especially their multimeric units often exceed this limit [42], as they tend to have relatively high mass to charge ratios [43–45], improvements were necessary in order to successfully transmit and measure these molecules in their native state, and also, ideally, in the form of complex, non-covalent adducts.

Douglas and French found in their work published in 1991 [46] that there is a correlation between applied pressure for ion trajectory optimization and ion mass when using a radio-frequency (rf) quadrupole. They found that somewhat counter-intuitively an increase in pressure led to improved ion transmission, similar to a process that had already been found for three-dimensional ion traps and was termed collisional damping or collisional focusing [46]. During the jet expansion of the ion stream moving from the capillary tip into the MS setup, some ions experience a deviation from their initial trajectory and once within the mass spectrometer, they acquire extra kinetic energy due to the m/z dependent acceleration via the applied voltages. If the spectrometer is not optimised for higher m/z values, these species might subsequently not be detected if they deviate too far from their initial trajectory. This can be prevented by reducing their initial kinetic energies [44]. Collisional damping does this by restricting the off-axis movements (with the axis being the main pathway of ions through the MS setup) and by reducing the speed of the ions both of which focus the ion beam. The movement is restricted by collisions with background gas atoms which, by the same process, also activate the ions by increasing their internal energy [14, 44] leading to the loss of attached solvent molecules.

It was found that reducing the frequency of the rf quadrupole led to a wider m/z coverage reaching up to values of 8585 m/z [47] which was further improved by Sobott *et al.* by decreasing the frequency even further and controlling the pressure at every step within the instrumental MS setup leading to the first use of a so-called tandem MS approach to analyse large assemblies of biomolecules [41].

Aside from the optimization of the quadrupole usage, the invention of nESI also greatly facilitated the transmission of high molecular weight ions as droplet sizes were decreased and desolvation conditions changed to be more 'gentle' (e. g. a reduced drying nitrogen flow rate in the ESI interface) which allowed non-covalent complexes to stay intact during measurements [48].

Quadrupole mass analysers

Quadrupole mass analysers have the capability to select ions of different masses by generating an alternating electric field which makes use of the stability of the ion trajectories. Quadrupoles became popular due to their low cost and modest power requirements as well as their relatively small size [49, 50] and were commercialised soon after their development [51]. They are often used at an early stage within a MS setup followed by further ion manipulation and a second, high-resolution mass analyser which is the final stage of ion detection.

Quadrupoles consist of four parallel circular or more often hyperbolic rods within which a potential distribution (rf voltage V) in the two directions perpendicular to the ion beam (X, Y) and an invariant potential distribution U along the ion beam or Z axis are applied. Ions travelling through the quadrupole and experiencing these electric fields can be modelled mathematically using Mathieu differential equations to determine the values of U and V for stable and unstable ion trajectories. An ion of a given polarity will move towards a rod of opposite polarity but if the potential of this rod changes sign before the ion discharges on contact, the flight direction will be changed. Repeating this several times for any given ion can stabilise its trajectory. If the magnitude of the applied voltages does not allow the trajectory of the ion to be changed in time, it will collide with the rod and not be measured.

Static voltages could shape the ion beam but varying the potential differences makes the mass analyser mass selective as it stabilises the trajectory of some ions while destabilising the path of others [49, 52], as displayed in Figure 2.2, which ultimately accounts for its m/z selectivity.

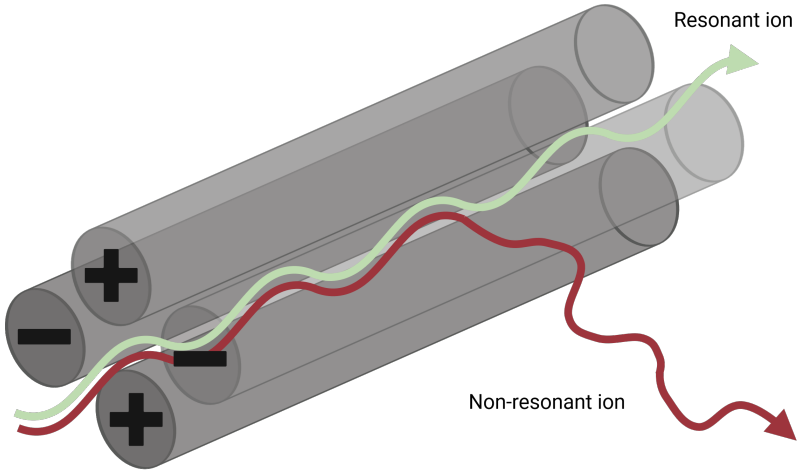

Figure 2.2: Quadrupole mass analyser. Schematic of a quadrupole mass analyser consisting of four rods with potentials of different signs at neighbouring rods. Trajectories of resonant and non-resonant ions are indicated. Figure created with BioRender.com.

Time-of-flight mass analysers

A time-of-flight (ToF) mass analyser or ToF-MS setup was first described by Wolff and Stephens in 1953 [53]. Ions reaching the mass analyser are accelerated in an electric field and then separated by their velocities while drifting through a field-free region called the flight tube. From the acceleration region ions are pulsed in discrete packages through the flight tube towards the detector. By this process all ions acquire the same kinetic energy and as a consequence their distribution of masses translates into a distribution of velocities. The acceleration of the ions by an applied potential V translates its electric potential energy E_{el} into kinetic energy E_{k} as described by equation 2.1 where m is the mass of an ion, $q = ze$ its charge and v its velocity [3, 52].

$$E_{\mathrm{k}} = \frac{mv^2}{2} = qV = zeV = E_{\mathrm{el}} \tag{2.1}$$

Solving equation 2.1 for the velocity v and substituting it into equation 2.2

$$t = \frac{L}{v} \tag{2.2}$$

where t is the time of flight it takes to cover the path length L, yields equation 2.3

$$t = \frac{L}{\sqrt{2Ve}}\sqrt{\frac{m}{z}} \tag{2.3}$$

which, given that L, V, and e are all constants, clearly shows that the time of flight of any given ion is directly dependent on its mass to charge ratio and can therefore be calculated when the time it takes until it hits the detector is known. Additionally, the equation shows that the lighter the ion, the faster it will reach the detector.

ToF mass analysers are popular as their design and usage are relatively easy, they produce a fairly high mass resolution and are theoretically capable of measuring across an unlimited mass scale [54]. However, if the flight tube is linear, which was usually the case during early ToF instrument development [55], the range of translational kinetic energies that exists for ions of a given m/z value could lower the resolving power of the instrument. To compensate for this, so-called reflectrons were incorporated into the setups. First developed by Mamyrin *et al.* and published in 1973 [56], these consist of a stack of ring electrodes with the opposite polarity to the ions with a decreasingly attractive potential. They are also called ion mirrors as they decelerate penetrating ions during their path in the flight tube and reflect them towards the detector as displayed in Figure 2.3A. Since ions with a larger kinetic energy penetrate further into the stack or electrodes, their reflection takes longer which focuses the ion beam of a particular m/z value. This narrows the peaks in the mass spectrum which improves the resolving power of the setup [3].

Another problem with the early instruments was the efficient interface of a continuous ion beam such as created by an ESI application with a ToF mass analyser that requires pulsed ion sampling in order to be able to measure flight times. This led to the advent of orthogonal acceleration ToF instruments where the ToF is situated in the orthogonal direction to the ion beam. Consequently, the pusher at the entrance to the ToF applies an orthogonal force to the ions which

they experience equally and independently of their velocity [57]. Furthermore, scattering of ions due to high pressure regions within the instrument can decrease the sensitivity at pressure values as low as $7 * 10^{-7} - 13 * 10^{-7}$ mbar which is why these instruments usually need to be operated at relatively low pressures to prevent broadening of arrival times [58].

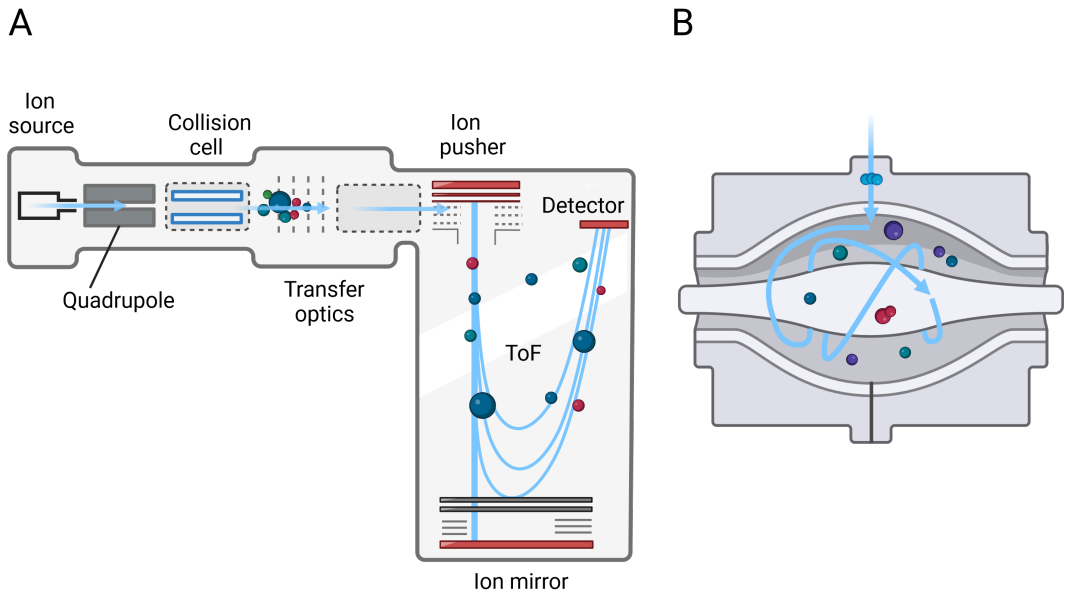

Figure 2.3: Quadrupole time-of-flight and orbitrap mass analysers. A Simplified schematic of an orthogonal quadrupole time-of-flight mass analyser with all the important components indicated. **B** Schematic of an orbitrap mass analyer. Figure created with BioRender.com.

Orbitrap mass analysers

Orbitrap analysers are electrostatic traps consisting of a barrel-shaped outer pair of electrodes which are usually at ground potential between which the ions are injected (see Figure 2.3B) as well as a spindle electrode in the middle between the outer electrodes which, for positive ions, is at a negative polarity of several kilovolts. Exclusively direct currents (DC) are applied within this mass analyser. Ions are usually injected with a kinetic energy of several keV and start to rotate in spirals around the inner electrode while oscillating harmonically along its axis as displayed in Figure 2.3B [52, 59]. The frequency of oscillations depends on the

ion's m/z value. When the ions come close to the outer electrodes this induces an image current. These current transients can be measured and depending on the frequency of the harmonic oscillations of the ions around the inner electrode, the different m/z values can be measured after Fourier transformation [60]. The injection time of ions of the same m/z values is shorter than the oscillation time around the inner electrode which is what allows the ions to oscillate coherently and enables measurement of distinct m/z values.

The electrostatic field in which the ions are trapped can be described by

$$U(r,z) = \frac{k}{2}\left(z^2 - \frac{r^2}{2}\right) + \frac{k}{2}(R_m)^2 ln\left(\frac{r}{R_m}\right) + C \tag{2.4}$$

where r and z are cylindrical coordinates with $z = 0$ being the axis of symmetry in the field, k is the field curvature, R_m is the characteristic radius, and C is a constant.

The orbitrap was first proposed by Makarov and published in patents towards the end of the 20th century [60–62]. It is a product of the search for instruments with a better ion transmission capacity and a greater intensity range for more accurate mass detection than Q-ToF instruments were able to offer [59]. In 2018 Fort *et al.* reported an expansion of instrument capabilities leading from the detection limit of 20,000 m/z to the measurement of ion signals with up to 70,000 m/z [63] with increased sensitivity and transmission properties that would allow regular use of orbitraps for high molecular weight biomolecules.

The collision cell

The collision cell is an important part of what is called tandem MS or MS/MS, the process of coupling two or more mass analysers to each other within a MS setup to increase the amount of structural information. For the collision cell to aid in this process, ions of a certain m/z range are first isolated by the quadrupole and then the chosen subset is submitted into the collision cell, which is a chamber full of neutral gas molecules. When ions are activated in the gas phase by being accelerated in such a cell, they undergo collisions with the neutral gas molecules during which a part of their kinetic energy is translated into internal energy and

another part transferred onto the neutral gas molecules. Depending on the voltage gradients and pressures applied to the cell, these collisions can lead to fragmentation of the monitored biomolecules or, in case they exist as a complex, to the dissociation of subunits whereby this usually leads to an unfolded, highly charged monomer and the complex of the remaining subunits [64, 65]. This method is also used for membrane dissociation of complexes or proteins which are embedded in membranes [66]. The whole process is called collision-induced dissociation (CID) or, in the case of orbitrap instruments, higher energy collisional dissociation (HCD) [67].

2.1.2 Solution hydrogen/deuterium exchange mass spectrometry

Hydrogen/deuterium exchange mass spectrometry (HDX-MS) is a special type of mass spectrometry that uses all the standard components described above but is coupled to a setup that allows one to investigate conformational changes and dynamics of proteins and their interaction partners by monitoring solvent exchange over time. This allows monitoring of local unfolding or dynamic opening and closing events as well as the (un)covering of interaction surfaces in a time dependent manner [68].

Aside from HDX-MS, the HDX methodology can also be used coupled to nuclear magnetic resonance (NMR) setups. Historically, this method was actually employed first before the advent of ESI made it possible to couple HDX to MS [69]. HDX-NMR is particularly helpful for the elucidation of protein folding at the residue level [70], but HDX-MS has the unique capability to investigate protein dynamics by elucidating interaction surfaces, conformational changes and analysing increasingly more complex mixtures [69] and was the focus of this work. Nowadays, both methods are often used to complement each other.

HDX-MS can be performed in the gas-phase where the H-D exchange takes place inside the spectrometer followed by measuring the intact mass to get information on the global deuterium incorporation [71]. However, this section will focus on solution HDX-MS as this was the method of choice for the present work in order

to access conformational changes across the whole amino acid sequence making up the proteins under investigation.

The standard 'bottom-up' procedure of a solution HDX-MS experiment is displayed in Figure 2.4. A buffer containing a protein or a mixture of proteins, is subjected to an excess of deuterated water, so that over time the hydrogen atoms ($\sim$1.0080 Da) of the surface accessible areas of the protein will be exchanged for heavier deuterium atoms (2.0141 Da) [68]. The longer the exposure to deuterated water, also called the labelling time, the higher the uptake of the surface exposed areas, depending on the ease of solvent accessibility. Usually, a set of several time points will be chosen for investigation for each experiment. To stop the labelling procedure, the solution is quenched by adding a buffer that brings it down to 0°C and pH 2.5. The quench buffer usually contains non-deuterated water as this is significantly cheaper than deuterated water. It also often contains chaotropic agents such as guanidinium hydrochloride (GdnHCl) to denature the proteins and aid downstream proteolysis. The low temperature minimises the kinetic energy of the proteins and therefore the back-exchange rate from deuterium to hydrogen. The low pH contributes to minimising the back-exchange rate [72, 73].

After quenching the samples can either be flash-frozen in liquid nitrogen and stored at 80°C until further use or immediately injected into the HDX manager which is a chamber that is kept at 0°C for the duration of the experiment. The proteins first reach a pepsin column where they get digested into peptides. This column is usually kept at low temperatures but above 0°C to ensure complete digestion of the protein. Pepsin is usually chosen for the digestion as it tolerates low pH values and is less specific which offers a redundancy in sequence coverage consisting of overlapping peptides of slightly different lengths increasing the certainty of observed mass increases. After the digestion step the peptides reach the desalting column which, as the name suggests, cleans up the peptides, followed by a separation step on a reversed-phase column which elutes the peptides through a gradient of increasing acetonitrile into the mass spectrometer which measures their m/z values and abundance over time. The time spent on the columns does not usually exceed

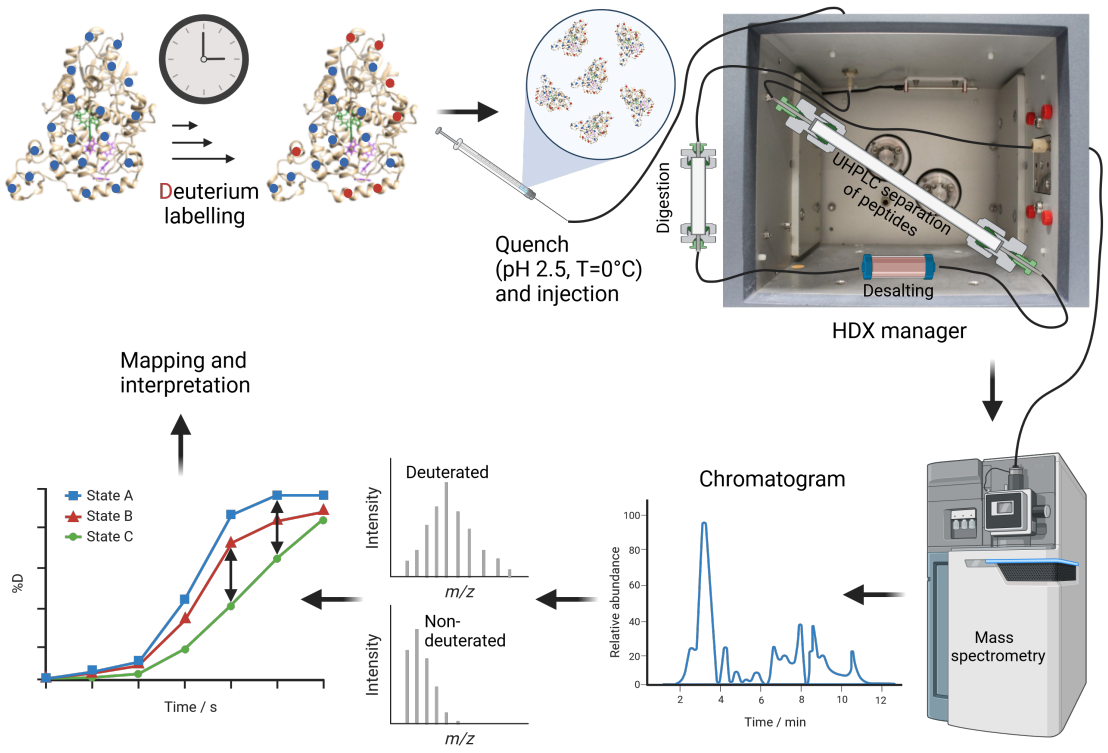

Figure 2.4: The common 'bottom-up' HDX-MS workflow. Proteins are incubated in excess deuterated water for different durations. The deuteration reaction whereby the solvent accessible protons get exchanged for deuterium is subsequently quenched by adding a buffer that lowers the pH to 2.5 and the temperature to 0°C followed by immediate injection into the HDX manager kept at 0°C. The proteins first undergo digestion in a pepsin column at 15-20°C. Afterwards, the peptides run at 0°C through a desalting column and a reversed-phase analytical column for separation. A mass spectrometer measures the m/z values over time and expresses their intensities over time as a chromatogram. Spectral analysis yields the isotope distribution envelopes for each peptide at each time point for comparison with a non-deuterated control protein run. The deuterium uptake can be displayed as uptake curves over time which can then be used for final interpretation and mapping of the uptake differences onto three-dimensional representations of the protein. Figure created with BioRender.com.

15 min. Usually, a ToF instrument is used for this purpose instead of trapping instruments as it has been argued that the latter could sometimes distort the isotope patterns [74]. In case of poor peptide separation due to the required low temperatures and steep and fast elution gradients, either the elution chromatography can be improved or ion mobility spectrometry can be added [68]. For the present work the peptides were sufficiently well separated and therefore ion mobility was not applied. Once the proteins are separated and eluted, they are subjected to electrospray ionisation (ESI) [13, 75–77].

After data acquisition, the spectral analysis is performed either by commercial softwares such as ProteinLynx Global Server (PGLS) by Waters, U.K., or non-commercial software such as the license-free software MASCOT [78]. The results are isotopic envelopes for each peptide at each measured charge state with the deuterated spectra moving up the m/z scale with respect to the non-deuterated controls. The controls have not been subjected to any prior deuterium labelling but instead to a non-deuterated equilibrium buffer and are a vital part of the data analysis process as they allow to measure the extent of the deuterium incorporation. These data sets usually require manual curation within a designated analysis program such as the commercially available DynamX 3.0 software (Waters, U.K.) or other, non-commerically available tools. A list of these has been compiled by Claesen *et al.* [79]. After curation, the data can then be expressed as percentage uptake plots over time. If HDX-MS runs are being performed for a control and a perturbed set of proteins, e.g. the same protein in different environmental conditions, these states can be subtracted from each other and their statistical relevance analysed. These differences, which are indicated with two-headed black arrows in Figure 2.4 at the bottom left, can subsequently be mapped onto three-dimensional protein structures obtained from X-ray crystallography, cryo-electron microscopy, or molecular modelling to facilitate the analysis of which regions experienced a change as a result of the investigated perturbation.

The uptake measured is a function of deuterium incorporation into the backbone of the proteins [80, 81] under investigation as the hydrogens of the side chains

exchange too quickly for the exchange to be quenched. Therefore, only the labile and surface-accessible backbone amide hydrogens exchange for deuterium and (mostly) stay exchanged under quenching conditions [1].However, H/D scrambling might still occur which is generally associated with slow heating conditions which the proteins encounter in between quenching and detection. It could obscure the data which is why over time fast-electron dissociation techniques such as electron-transfer dissociation (ETD) and electron-capture dissociation (ECD) [82] were explored as a means to overcome these technical difficulties [83]. However, despite having advantages such as better spatial resolution [84], these techniques are still challenging to employ. Another, more recently developed technique is ultra-violet photo dissociation (UVPD), a method that increases the energy of the ions by high-energy photon absorption leading to the exploration of different dissociation pathways [85–87].

Aside from the 'bottom-up' approach described above which was used for the current work, 'top-down' and 'middle-down' approaches exist as well. For the 'top-down' approach, the quench does not denature the protein. Although some fragmentation steps might be applied within the mass spectrometer, the final measurement will be of whole proteins or polypeptides without prior digestion, cleaning or separation steps. For the 'middle-down' approach, the proteins are also not digested, but desalted and separated on a reverse phase column, followed by ETD or ECD which enable analysis of the mass increase at a peptide or even single amino acid level.

Another problem that HDX-MS experiments often suffer from is carryover. This is measuring a mixture of a protein and some impurities in the form of peptides from a previous measurement that were not fully eluted from the lines and columns. Several protocols for the reduction of carryover have been used and described in the literature [72] of which the injection of a chaotropic agent after each sample injection was chosen for the present work.

The labelling time points acquired during an HDX-MS experiment are usually limited towards the lower end to several seconds by either the accuracy with which a

liquid handling robot can perform the labelling tasks or by how fast an experimenter can move. However, ways to circumvent this problem have been explored with labelling for very short times on ice instead of at room temperature to slow down the kinetics of the exchange reaction [88, 89], by using subsecond microfluidics systems coupled to HDX-MS [90], setups that reach millisecond time resolution [91, 92] which includes the 'air gap' method that can be applied using a conventional liquid-handling robot such as the LEAP HTS PAL [93], as well as the use of liquid droplets that even reach submillisecond resolution [94].

The equilibration or continuous labelling technique described above is commonly employed when using HDX-MS as a tool for protein investigation and was also used for the work described in this thesis. However, it is also possible to apply pulsed labelling techniques in which the labelling times are short and of a fixed length. This approach can allow insight into the folding processes of proteins providing snapshots during the process and reveal which parts of the structure are disordered or unfolded [95, 96].

HDX-MS kinetics

The rate of H/D exchange is influenced by numerous factors such as pH and temperature as mentioned above, but also by neighbouring amino acids [97], pressure, and ionic strength [95, 98]. The rate constant for a given backbone amide hydrogen exchange is given by

$$k_{\mathrm{ch}} = k_{\mathrm{int,acid}}[\mathrm{H_3O^+}] + k_{\mathrm{int,base}}[\mathrm{OH^-}] + k_{\mathrm{int,water}}[\mathrm{H_2O}] \tag{2.5}$$

which describes both the acid catalysed and base catalysed version of the solvent exchange where k_{ch} is the total intrinsic chemical rate of HDX, and $k_{\mathrm{int,acid}}$, $k_{\mathrm{int,base}}$, and $k_{\mathrm{int,water}}$ are the rates for the acid, base and water catalysed reactions, respectively [1, 68, 99].

Peptide backbones can be in an 'open' and solvent accessible state and in a 'closed' and inaccessible state. The open state is considered to be the state in which the backbone is sufficiently accessible for the formation of hydrogen

bonds between the amide hydrogen and a water molecule in the solvent so that an exchange of atoms is possible. As the two states can constantly interchange, this can be described as follows

$$\mathrm{N-H_{closed}} \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} \mathrm{N-H_{open}} \xrightarrow{k_{ch}} \mathrm{N-D_{open}} \underset{k_{op}}{\overset{k_{cl}}{\rightleftharpoons}} \mathrm{N-D_{closed}} \tag{2.6}$$

where k_{op}, k_{cl}, and k_{ch} correspond to the opening, the closing and the backbone exchange rate constant as defined in equation 2.5 above, respectively. Since the proteins will be subjected to an excess of deuterated water, [D]$\gg$[H] is assumed and therefore the rate of exchange, k_{ex}, can be expressed as

$$k_{ex} = \frac{k_{op}k_{ch}}{k_{cl} + k_{ch}} \tag{2.7}$$

if $k_{ch} \gg k_{cl}$, it would follow that $k_{ex} \approx k_{op}$ which means that during each opening event a complete exchange reaction takes place. This leads to two distinct distributions, one where slow opening events have been labelled and one where the opening did not take place during the labelling time, leading to a bimodal isotopic distribution in the mass spectrum. If on the other hand, $k_{ch} << k_{cl}$, the exchange reaction cannot take place completely and only a part of the transiently accessible area is labelled. The corresponding isotopic envelopes will display a shift towards higher m/z values over time. These kinetic regimes are known as EX1 behaviour for the former and EX2 behaviour for the latter and illustrated in Figure 2.5. EX2 is assumed to be the more common kinetic behaviour with flexible and dynamic regions switching rapidly between the opened and closed position and more rigidly structured regions undergoing these types of transitions once within a time frame of several minutes or even days. EX1 might be observed if large restructuring events take place during the labelling time [68]. However, EX1 might not capture the full range of possible protein dynamics as it does not consider intermediate movements that are not captured by a binary interpretation of events [100]. It is also possible to observe mixed EX1/EX2 behaviour [101, 102].

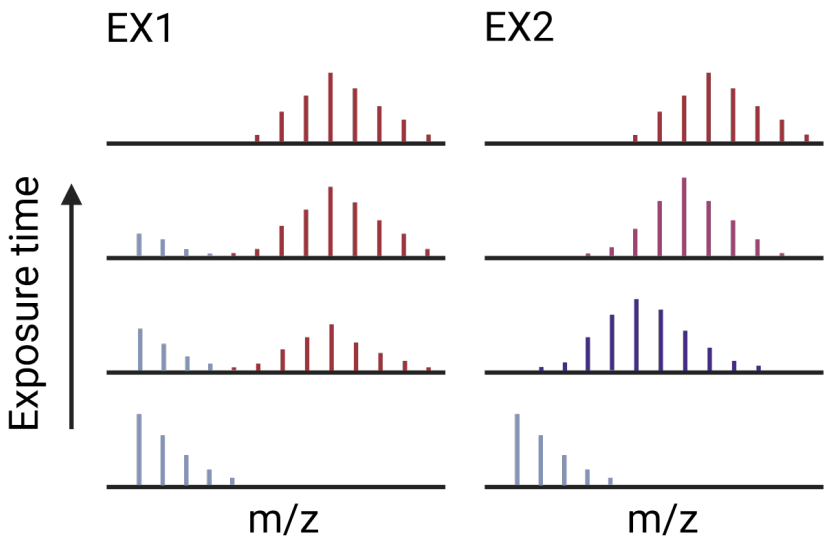

Figure 2.5: HDX-MS kinetics: EX1 & EX2. HDX-MS spectra of any given peptide as a function of exposure time. Left: EX1 kinetics showing two different states that exist in parallel, with one 'open' and therefore easily accessible isotopic distribution and a 'closed' and therefore less accessible one. Right: EX2 kinetics of one peptide displaying a gradual increase in deuteration towards higher m/z values. EX2 behaviour is a more common pattern observed for native proteins. Figure created with BioRender.com.

2.1.3 Mass photometry

Mass photometry (MP) is a novel tool for bioanalytical analysis. Similar to mass spectrometry, it measures masses of biomolecules such as proteins. It yields rapid results and can detect sample heterogeneity, differentiate between monomeric compounds and complexes of biomolecules, and it works in solution at room temperature without the need for labelling. In contrast to mass spectrometry, which ionises molecules before measuring their m/z values, MP uses light scattering which directly scales with the mass of the biomolecules causing the scattering. An advantage of this technique in comparison to mass spectrometry is that it only requires extremely small sample volumes in the nanomolar range. Additionally, samples do not require modification; it is possible to measure them in their original buffer which is usually not the case for mass spectrometry. MP can measure molecules within a mass range of 30 kDa - 5 MDa [103, 104] with a 2% mass accuracy and a resolution of up to 19 kDa (FWHM) [104–106] which is where it

falls a little behind the resolving capabilities of mass spectrometry.

Mass photometry was first introduced in 2018 by Young *et al.* [105]. Its working principle is based on using interferometric scattering microscopy [107] to detect light scattering of nanoscopic objects where the contrast resulting from the scattering depends on the differences in refractive indices of the measured biomolecules and the surrounding medium. As proteins all have similar optical properties and densities, the contrast is related to the protein mass in a proportional manner [106].

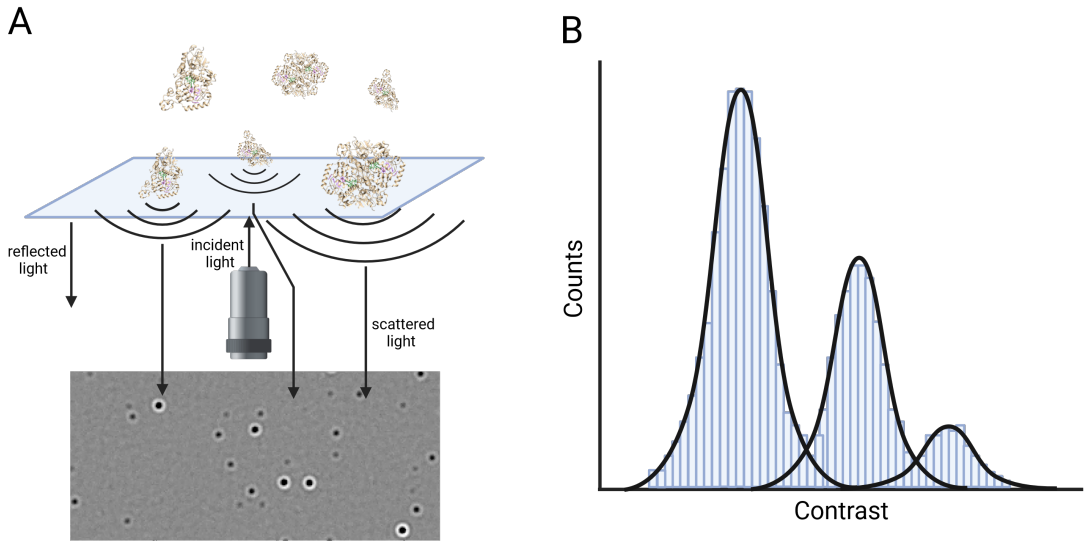

Figure 2.6: Working principle of mass photometry. A Binding events cause local changes in the scattering of the incident light. The events are recorded as ratiometric images and perceived as intensity changes which are the direct result of changes in reflectivity. The contrast scales with the mass of the protein or protein complex. **B** Contrast histogram displaying the distribution of species in solution. Each species is usually fitted with a Gaussian. Figure created with BioRender.com.

To perform a mass photometry measurement, a drop of a solution containing the analytes is placed on a cleaned glass microscope slide and illuminated by a laser light which reaches the glass-water surface and is reflected by it. Once a biomolecule adsorbs onto the glass microscope slide, it influences the scattering of the incident light as displayed in Figure 2.6A. The scattering is recorded as ratiometric images which are a rolling average of natives frames. The relative difference of two averages yields the visual results displayed at the bottom of

Figure 2.6A. All molecules detected during the time frame of an experiment can be combined and displayed as a histogram of counts as a function of contrast as in Figure 2.6B. This histogram represents the distribution of species in solution and can be translated into their masses aided by a previously performed contrast measurement of a range of standard biomolecules [104, 105].

2.2 Electron Paramagnetic Resonance Spectroscopy

Electron paramagnetic resonance spectroscopy (EPR), also known as electron spin resonance spectroscopy (ESR), is a method that studies systems with unpaired electrons and their resonant absorption of microwave radiation. EPR is a popular tool used in various scientific fields such as materials science [108], environmental chemistry [109], and the study of free radicals [110]. It is also used for medical research, e.g. to determine tissue oxygen and radiation dose in clinical settings [111] as well as to investigate aspects of structural biology, often by attaching labels containing unpaired electrons to specific protein regions to obtain distance distribution information [112].

In EPR, the magnetic component of the electromagnetic radiation interacts with the permanent magnetic dipole moments of the unpaired electrons and an externally applied magnetic field aligns the dipoles. The interaction between the unpaired electrons and the magnetic field is called the *Zeeman* interaction and its resulting effect is a splitting of energy levels where the electron magnetic moments are either parallel or anti-parallel to the applied field, that is they exist in different spin sub-levels. Spin is a quantum mechanical property describing the intrinsic angular momentum of an elementary particle that can exist in different quantised states [113]. Resonant absorption, which leads to a transition between the energy levels, occurs when the energy gap between two energy levels matches the electromagnetic frequency of the microwave radiation. In the absence of hyperfine interactions, this is described by the following equation

$$\Delta E = h\nu = g\mu_{\mathrm{B}}B \tag{2.8}$$

where E is the energy, h is the Planck constant, ν the frequency of the microwave radiation, g is the g-factor which is a correction or proportionality factor dependent on the local environment of the unpaired electron and which corrects for deviations from the free electron g-value (2.0023), μ_B is the Bohr magneton, and B is the magnitude of the externally applied magnetic field. This resonant absorption is what is measured in a continuous wave EPR (cw-EPR) spectrometer [113]. The most commonly employed microwave frequencies are 9.5 GHz, 34 GHz, and 95 GHz, labelled as X-, Q-, and W-band, respectively. The main components of any EPR spectrometer are a microwave radiation source, a magnet to apply a homogeneous magnetic field, a resonator cavity in which the incoming microwave radiation is concentrated on a sample and the output amplified, a cryo- or thermostat to control the temperature of the sample, and a control and signal processing unit [114]. There are different modes of EPR spectroscopy and in the following a brief introduction to continuous wave and pulse EPR - with a focus on double electron-electron resonance (DEER) - will be given. The introductions in this work will be brief as the main focus of this thesis is the use of mass spectrometry techniques.

2.2.1 Continuous wave EPR

In cw-EPR, the microwave radiation is kept at a fixed frequency and amplitude and applied continuously to a sample while an externally applied magnetic field is swept. Once the magnetic field induces a splitting of the energy levels fulfilling the resonance condition in equation 2.8, absorption of the microwave radiation occurs inducing transitions between energy levels.

The microwave radiation reaching the resonator cavity forms a standing wave. The cavity is usually designed in a way that maximises the magnetic component of the radiation in the middle of the cavity and equally minimises the electric component at the same position. For optimal results, a tuning procedure is carried out by fine-tuning the applied frequency and by increasing or decreasing the opening to the cavity. This fine-tunes the coupling of the radiation into the cavity and minimises the reflection from the cavity. Absorption of microwave radiation by

the sample in the cavity results in a change in impedance of the system and a corresponding increase in reflected microwave power which is detected.

As the changes in impedance due to absorption can be very small, the noise of the measurements gets reduced by applying a weak additional magnetic field at a modulation frequency of approximately 100 kHz via Helmholtz coils exterior to the cavity. Phase-sensitive lock-in detection extracts solely the absorption signal at the modulation frequency. As a consequence of this process, the cw-EPR signal is obtained as the first derivative of the microwave absorption [113, 115].

2.2.2 Pulse EPR

Pulse EPR was invented after cw-EPR to overcome its spectral and time limitations. This technique enables to assess the different parameters of a spin system by separating different interactions from each other. In pulse EPR, instead of continuously irradiating with microwaves, high intensity microwave pulses are applied to a sample; these pulses are usually of the order of nanoseconds. Different techniques have been developed to measure different properties of the systems under investigation. All of these techniques consist of the same basic components described above but are modified with a different, specifically designed resonator and by applying additional pulse forming units.

The short, high power microwave pulses applied in pulse EPR are used to manipulate the magnetisation of the sample. Usually, $\frac{\pi}{2}$ and π pulses are applied in different sequences, turning the bulk magnetisation of the sample from the z-direction (which coincides with the strong static magnetic field direction) into the xy-plane (a $\frac{\pi}{2}$ pulse) or onto the -z direction (a π pulse). Therefore, changes in spin behaviour or their precessing frequencies are usually detected in the xy-plane. Observation of the time evolution of the magnetisation in the xy-plane following the $\frac{\pi}{2}$ pulse results in the detection of a so-called free induction decay. However, due to instrumental protection mechanisms overlapping with the signal, these are often not measured directly. A second pulse is applied following an inter-pulse delay τ, leading to a refocusing of the magnetisation at time 2τ after the initial pulse.

The pulse EPR signal is then acquired via detection of this refocused magnetisation, also known as the spin echo [116].

The pulse EPR technique used for the work in this thesis is double electron-electron resonance (DEER), a technique which investigates the dipolar coupling between two electron spins using microwave pulses at two different frequencies with the aim to measure the inter-spin distances.

DEER

DEER, also known as pulsed electron double resonance (PELDOR), uses two separate microwave frequencies to individually excite two subsets of spins. Through measurement of the inter-electron spin dipolar interaction, distances between radicals can be determined. The spins are often nitroxide spin labels which are artificially introduced radicals which, in the case of proteins, can be introduced in a site-directed manner, enabling distance measurements between different parts of a protein or between different proteins [117].

The distance r between two spin centres A and B can be calculated with equation 2.9, which is an approximation neglecting orientational effects,

$$r = (g_A g_B \mu_B^2 \frac{\mu_0}{4\pi\hbar\omega_{dd}})^{1/3} \tag{2.9}$$

where g_A and g_B are the g-factors of the two spin centres, μ_B is the Bohr magneton, μ_0 the magnetic permeability of vacuum, $\hbar$ the reduced Planck's constant, and ω_{dd} the dipolar frequency which can be determined from the time-trace recorded in DEER experiments [114].

To generate and record these time traces, one spin ensemble is chosen to be the 'detector' spins and another ensemble comprises the 'pump' spins. Even if both spins originate from the same type of molecule (e.g. two identical nitroxide labels), differentiation is possible due to the capability of the method to pump on different transition manifolds, aided by the spins' different local environments and orientations [118]. The pulse sequence $\frac{\pi}{2}$ - τ_1 - π - τ_1 - echo is applied to the detector spins which refocuses the spin magnetisation to produce a measurable spin echo,

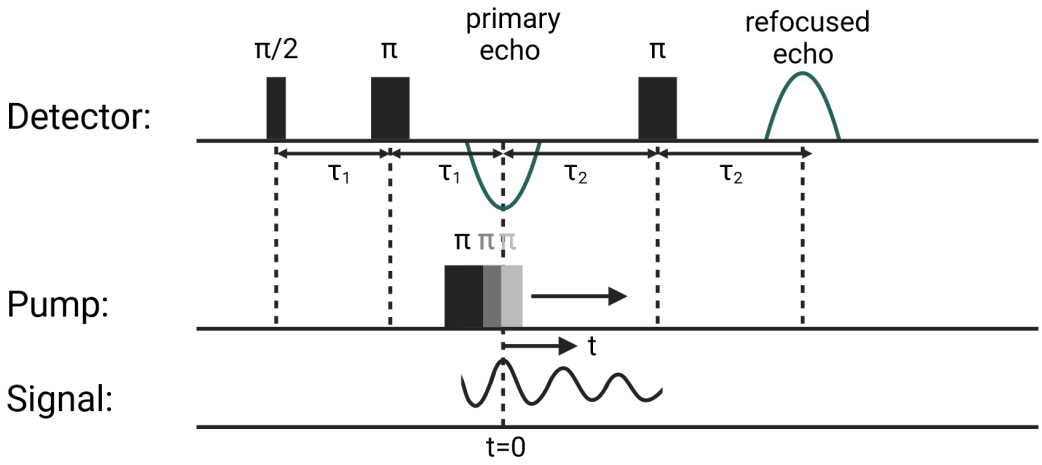

Figure 2.7: Pulse sequence for 4-pulse DEER. Detector and pump DEER sequence with corresponding signal. The indicated time axis corresponds to the time position of the pump pulse with respect to the primary echo and the signal is monitored as a function of the position of this pump pulse in time with one data point corresponding to one pump pulse position. At the zero time the first refocused echo is measured. Figure created with BioRender.com.

also called Hahn echo. In this work, the 4-pulse version of DEER displayed in Figure 2.7 was used. It avoids the zero-time artefacts occurring in 3-pulse DEER (which was used more commonly before the 4-pulse sequence was developed) due to an overlap of the detection and pump pulses. In 4-pulse DEER a 3-pulse refocused-echo sequence forms two spin echos. Detection is only performed on the second echo. After an initial two pulses, a $\frac{\pi}{2}$ followed by a π pulse, another π pulse at a different microwave frequency inverting the pump spin populations is applied which is then varied in time up until a second π pulse is applied to the detector spin population to refocus the spin magnetisation and obtain the second spin echo. The intensity of this refocused echo is therefore monitored in time as a function of the position of the pump pulse and is dependent on the effect the pump pulse had on the magnetisation of a subset of spins. The oscillating frequency of the refocused echo is a measure of the strength of the coupling between the two spin ensembles and can be directly related to the distance between them [112, 119].

2.3 Molecular Dynamics Simulations

Molecular Dynamics (MD) simulations are a tool to visualise and calculate the motions of a biomolecular system. It is an atomistic method which analyses the movements of the system at the level of atoms. This is done by solving the simultaneous classical equations of motions for hundreds of atoms at the same time using Newton's laws of motion. By calculating the force exerted on any given atom by all the surrounding atoms, the position of every atom can be calculated. If this process is repeated, the atoms' positions and velocities can be followed over time [120].

MD simulations have become increasingly popular over the years starting with the first MD simulation of gases in 1957 [121] followed by the first protein simulation by McCammon *et al.* of the bovine pancreatic trypsin inhibitor in 1977 [122]. They have gained popularity due to increased computing power allowing simulations on longer time-scales while simultaneously reducing the computing time [123]. Especially in the field of neuroscience this method is extensively used, aiding for example in the understanding of aggregation processes associated with neurodegenerative diseases or providing additional information about drugs targeting the nervous system. However, also other fields concerned with the structural or functional analysis of biomolecules have joined the trend to use MD to 'fill in the gaps', aid with the interpretation of experimental data, and to guide the design of new experiments [120].

The above described classical approach of MD takes into account steric hindrance and electrostatic interactions described by a force field, but only quantum mechanical approaches offer the integration of bond cleavages, charge transfers and electronic excitations which would in principle be an interesting feature to add for the present work as the radical pair mechanism includes excited states and charge transfers. However, these methods are generally restricted to a few hundred atoms due to their immense computational cost. In 1976 Warshel and Levitt [124] widened the capabilities of quantum mechanics (QM) with respect to large biomolecules by combining it with molecular mechanics (MM), its classical counterpart, which is now generally known as QM/MM. This approach allows a certain area of a

biomolecule to be calculated more accurately using quantum mechanics while the interactions of the regions of the protein requiring a less accurate approach are simulated using molecular mechanics. In principle, QM/MM can then be used within any MD method to simulate the interactions and movements of the atoms over time. QM/MM methods can combine different approaches of QM and MM, but they have the additional difficulty that special precautions have to be taken to efficiently couple the two systems to one another so as not to destroy or deform the environment within a biomolecule [125]. Therefore, although this approach has gained popularity, it requires more in-depth knowledge on the part of the user and often a simple classical MD approach might be sufficient to answer questions about the behaviour and interactions of a certain biomolecule.

MD has the additional benefit that solvent interactions can easily be integrated into the simulations as usually all biomolecules are immersed into a mixture of simulated water molecules and salt ions at varying concentrations whereas QM approaches often treat the molecules at low pressures and in the gas phase.

Aside from MD, the Monte Carlo (MC) approach is another method with similar capabilities [126]. In MC a random perturbation is introduced to the atom coordinates. If the overall energy decreases during this process, the movements are accepted and with a probability of $e^{-\Delta E/kT}$ they are also accepted if the overall energy increases. This allows sampling of a wider energy surface, exploring local maxima and local as well as global minima. The geometry changes can be caused by 'non-physical' perturbations giving this method increased sampling capabilities in comparison to MD as configurations that would otherwise only be reached by overcoming a high energy barrier could easily be sampled via this approach [127]. This method does not calculate the atom velocities, only randomly moves one particle each step followed by a calculation of the change in energy. Therefore, the next step does not depend to the same extent on the previous step as is the case for MD simulations. For the purposes of this work, the latter was chosen as the more suitable technique.

2.3.1 Force fields and water models

The function of force fields within an MD simulation is to describe the energy of a molecule as a differentiable function of its geometry while also integrating a set of static parameters that describe the interactions between particles [128]. To achieve this, MD makes use of the Born-Oppenheimer approximation which assumes that nuclear and electronic motions happen on different timescales and can therefore be treated separately, by only concentrating on the nuclear motions [129].

The force fields consist of functional forms that aim to describe the properties of the molecules divided up into bonded and non-bonded properties. The energy resulting from bonded interactions E_{bonded} can be described by equation 2.10

$$E_{\text{bonded}} = \sum_{\text{bonds}} K_b(b-b_0)^2 + \sum_{\text{angles}} K_\Theta(\Theta-\Theta_0)^2 + \sum_{\text{dihedrals}} K_\chi[1+cos(n\chi-\sigma)], \quad (2.10)$$

where the first term describes the stretching of the bonds with b representing the interatom distance, K_b the stiffness, and b_0 the equilibrium length of the bond. The second term describes the bending of the valence angles considering a chain of three atoms with Θ as the angle between two bond vectors, K_Θ as the stiffness and Θ_0 as the equilibrium geometry of the angle, and finally the third term represents the rotation of the dihedrals defined by a chain of four atoms with χ being the value of the dihedral angle, K_χ the energetic parameter describing the barrier heights, n the periodicity, and σ the phase. The cosine captures the periodicity of the rotation and is added to 1 so that $E_{bonded} \geq 0$. The dihedrals usually also contain another term which describes the so-called 'improper' dihedrals. These are used to maintain planarity or chirality at an atom with three other atoms attached.

The energy resulting from non-bonded properties, $E_{\text{nonbonded}}$, is described by equation 2.11

$$E_{\text{nonbonded}} = \sum_{\substack{\text{nonbonded} \\ \text{pairs ij}}} \left(\epsilon_{ij} \left[\left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^{12} - \left(\frac{R_{\text{min},ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{r_{ij}} \right), \quad (2.11)$$

which combines the description of the electrostatics, dispersion, and the Pauli exclusion principle considering the interactions between atoms in separate molecules

or atoms separated by three or more bonds within one molecule. The part of the equation written within the square brackets is the Lennard-Jones (LJ) equation, also referred to as the van der Waals term. It describes the attractive and soft repulsive forces, the Pauli-exclusion interactions, with very close inter-atom distances r_{ij} being repulsive, moderate distances being attractive and long distances leading to little or no interaction. As the distance between two atoms decreases, the second term dominates leading to favourable, progressively more negative energies leading to attraction and as the distance further decreases the first term dominates leading to increased energies and repulsion. ϵ_{ij} is chosen dependent on the types of atoms between which an interaction takes place, as is $R_{\mathrm{min},ij}$ which also depends on the distance at which the LJ energy is at a minimum. The second part of the equation is Coulomb's law with q_i and q_j describing the charges on the atoms under consideration. To calculate the total energy of the system, the bonded and nonbonded energies as well as any force field-specific terms need to be combined. Force field parameters are usually optimized by trying to model a target data set that can originate from experimental techniques such as vibrational spectra or X-ray diffraction structures [128, 129].

There is a wealth of different force fields available to the user. Currently, popular choices include AMBER [130], a group of force fields originally developed in the program of the same name, CHARMM [131], a force field of which also many different versions exist, or GROMOS [132], which is primarily coded within the program of the same name. Often, different force fields have been designed for specific purposes such as a specialisation to model protein movements [133]. Additionally, MD simulations need a number of other parameters to be defined by the user. The different methods and variables are discussed in section 2.3.2.

As most biomolecules are simulated in an environment as close as possible to their native environment, which in a lot of cases is water-based, the choice of the water model is important to consider as it needs to be compatible with the force field of choice. Over the years multiple different water models have emerged out of which the TIP3P, TIP4P and SPC models are especially popular. These models

have in common that they use point interaction sites on their nuclei with separate partial charges used to calculate the intermolecular Coulomb energies [134].

2.3.2 Constraints and conditions: periodic boundary conditions, cutoffs, thermo- and barostats

Further to choosing a force field and water model, another common feature of MD simulations that is frequently implemented are periodic boundary conditions (PBC). These periodically repeat the solvent environment that a biomolecule is simulated in and thereby prevent it from suddenly encountering a vacuum while it is moving around during the simulation. Alternatively, the biomolecule can be enclosed in a sphere of solvent with forces at the boundary of the sphere mimicking a larger solvent area. However, the PBC has been shown to more accurately reproduce experimental data. One possible downside of PBC is that sometimes long-range electrostatic interactions can be reproduced periodically and lead to artefacts [135, 136].

As the cutoff size after which distance interactions between atoms are no longer taken into consideration influences the stability of the simulated system [137], it is important to chose the cutoff values such that the stability of common secondary structures such as α-helices is not diminished. An option that is often made use of instead of a sharp cutoff value, is the particle mesh Ewald summation (PME) where the cutoff value describes the distance at which a transition from short-range to long-range interactions takes place [138].

The constraints applied to keep the system stable and most 'native-like' can lead to a change in energy in the simulated system resulting in a change in temperature. Therefore, thermostats to keep the temperature at a constant value are usually integrated into the simulation which scale the velocities of the atoms such that a predefined temperature is kept constant [128]. Sometimes it might also be necessary to use a barostat to keep the pressure of the system constant, e.g. if exact experimental conditions need to be replicated. Like the thermostat, the barostat uses algorithms to manipulate the velocities of the atoms in order to control the pressure [139].

References

1. Kaltashov, Igor A and Eyles, S. J. *Mass spectrometry in structural biology and biophysics: architecture, dynamics, and interaction of biomolecules* 2nd ed. (John Wiley & Sons, 2012).
2. Dempster, A. J. A New Method of Positive Ray Analysis. *Physical Review* **11,** 316–325. https://link.aps.org/doi/10.1103/PhysRev.18.415 (1918).
3. McCullagh, J. & Oldham, N. *Mass spectrometry* (Oxford University Press, 2019).
4. Leney, A. C. & Heck, A. J. R. Native Mass Spectrometry: What is in the Name? *Journal of the American Society for Mass Spectrometry* **28,** 5–13. https://pubs.acs.org/doi/10.1007/s13361-016-1545-3 (2017).
5. Lorenzen, K., Vannini, A., Cramer, P. & Heck, A. J. Structural Biology of RNA Polymerase III: Mass Spectrometry Elucidates Subcomplex Architecture. *Structure* **15,** 1237–1245 (2007).
6. Glish, G. L. & Vachet, R. W. The basics of mass spectrometry in the twenty-first century. *Nature Reviews Drug Discovery* **2,** 140–150. https://www.nature.com/articles/nrd1011 (2003).
7. Tamara, S., den Boer, M. A. & Heck, A. J. R. High-Resolution Native Mass Spectrometry. *Chemical Reviews* **122,** 7269–7326. https://pubs.acs.org/doi/10.1021/acs.chemrev.1c00212 (2022).
8. Heck, A. J. R. Native mass spectrometry: a bridge between interactomics and structural biology. *Nature Methods* **5,** 927–933. https://www.nature.com/articles/nmeth.1265 (2008).
9. Yamashita, M. & Fenn, J. B. Electrospray ion source. Another variation on the free-jet theme. *The Journal of Physical Chemistry* **88,** 4451–4459. https://pubs.acs.org/sharingguidelines%20https://pubs.aip.org/jcp/article/79/12/6043/87775/On-the-time-required-to-reach-fully-developed-flow (1984).
10. Aleksandrov, M. *et al.* Formation of beams of quasimolecular ions of peptides from solutions (Russian). *Bioorganicheskaya Khimiya* **11,** 700–704 (1985).
11. Meng, C. K., Mann, M. & Fenn, J. B. Of protons or proteins. *Zeitschrift für Physik D Atoms, Molecules and Clusters* **10,** 361–368. http://link.springer.com/10.1007/BF01384871 (1988).
12. Smith, R. D., Loo, J. A., Edmonds, C. G., Barinaga, C. J. & Udseth, H. R. New developments in biochemical mass spectrometry: electrospray ionization. *Analytical Chemistry* **62,** 882–899. https://pubs.acs.org/sharingguidelines%20https://pubs.acs.org/doi/abs/10.1021/ac00208a002 (1990).
13. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **246,** 64–71. https://www.science.org%20https://www.science.org/doi/10.1126/science.2675315 (1989).

14. Benesch, J. L. P., Ruotolo, B. T., Simmons, D. A. & Robinson, C. V. Protein Complexes in the Gas Phase: Technology for Structural Genomics and Proteomics. *Chemical Reviews* **107,** 3544–3567. https://pubs.acs.org/doi/10.1021/cr068289b (2007).

15. Mass Spectrometry of Protein Complexes: From Origins to Applications. *Annual Review of Physical Chemistry* **66,** 453–474. https://www.annualreviews.org/doi/10.1146/annurev-physchem-040214-121732 (2015).

16. Morgner, N. & Robinson, C. V. Massign: An Assignment Strategy for Maximizing Information from the Mass Spectra of Heterogeneous Protein Assemblies. *Analytical Chemistry* **84,** 2939–2948. https://pubs.acs.org/doi/10.1021/ac300056a (2012).

17. Emmett, M. R. & Caprioli, R. M. Micro-Electrospray Mass Spectrometry: Ultra-High-Sensitivity Analysis of Peptides and Proteins. *Journal of the American Society for Mass Spectrometry* **5,** 605–613. https://pubs.acs.org/sharingguidelines (1994).

18. Wilm, M. S. & Mann, M. Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *International Journal of Mass Spectrometry and Ion Processes* **136,** 167–180. https://linkinghub.elsevier.com/retrieve/pii/0168117694040249 (1994).

19. Gale, D. C. & Smith, R. D. Small volume and low flow-rate electrospray lonization mass spectrometry of aqueous samples. *Rapid Communications in Mass Spectrometry* **7,** 1017–1021. https://onlinelibrary.wiley.com/doi/10.1002/rcm.1290071111 (1993).

20. Juraschek, R., Dülcks, T. & Karas, M. Nanoelectrospray—More than just a minimized-flow electrospray ionization source. *Journal of the American Society for Mass Spectrometry* **10,** 300–308. https://pubs.acs.org/sharingguidelines%20https://pubs.acs.org/doi/10.1016/S1044-0305%2898%2900157-3 (1999).

21. Wilm, M. & Mann, M. Analytical Properties of the Nanoelectrospray Ion Source. *Analytical Chemistry* **68,** 1–8. https://pubs.acs.org/sharingguidelines%20https://pubs.acs.org/doi/10.1021/ac9509519 (1996).

22. Susa, A. C., Xia, Z. & Williams, E. R. Small Emitter Tips for Native Mass Spectrometry of Proteins and Protein Complexes from Nonvolatile Buffers That Mimic the Intracellular Environment. *Analytical Chemistry* **89,** 3116–3122. https://pubs.acs.org/doi/10.1021/acs.analchem.6b04897 (2017).

23. Van den Heuvel, R. H. & Heck, A. J. Native protein mass spectrometry: from intact oligomers to functional machineries. *Current Opinion in Chemical Biology* **8,** 519–526. https://linkinghub.elsevier.com/retrieve/pii/S1367593104001048 (2004).

24. Bernstein, S. L. & Bowers, M. T. in *Comprehensive Analytical Chemistry* 63–82 (2008).
https://linkinghub.elsevier.com/retrieve/pii/S0166526X08002031.

25. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionization mass spectrometry. *Journal of The Royal Society Interface* **9,** 801–816.
https://royalsocietypublishing.org/doi/10.1098/rsif.2011.0823 (2012).

26. Xia, Z. & Williams, E. R. Protein-Glass Surface Interactions and Ion Desalting in Electrospray Ionization with Submicron Emitters. *Journal of the American Society for Mass Spectrometry* **29,** 194–202.
https://pubs.acs.org/doi/10.1007/s13361-017-1825-6 (2018).

27. Kebarle, P. & Verkerk, U. H. Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrometry Reviews* **28,** 898–917.
https://onlinelibrary.wiley.com/doi/10.1002/mas.20247 (2009).

28. Hernández, H. & Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nature Protocols* **2,** 715–726. https://www.nature.com/articles/nprot.2007.73 (2007).

29. Rayleigh, L. XX. On the equilibrium of liquid conducting masses charged with electricity. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **14,** 184–186.
https://www.tandfonline.com/doi/full/10.1080/14786448208628425 (1882).

30. On the evaporation of small ions from charged droplets. *The Journal of Chemical Physics* **64,** 2287–2294 (1976).

31. Dole, M. *et al.* Molecular beams of macroions. *The Journal of Chemical Physics* **49,** 2240–2249 (1968).

32. Fernandez de la Mora, J. Electrospray ionization of large multiply charged species proceeds via Dole's charged residue mechanism. *Analytica Chimica Acta* **406,** 93–104.
https://linkinghub.elsevier.com/retrieve/pii/S0003267099006017 (2000).

33. Pimlott, D. J. & Konermann, L. Using covalent modifications to distinguish protein electrospray mechanisms: Charged residue model (CRM) vs. chain ejection model (CEM). *International Journal of Mass Spectrometry* **469,** 116678.
https://doi.org/10.1016/j.ijms.2021.116678%20https://linkinghub.elsevier.com/retrieve/pii/S1387380621001585 (2021).

34. Fernández de la Mora, J. The fluid dynamics of Taylor Cones. *Annual Review of Fluid Mechanics* **39,** 217–243 (2007).

35. Sharon, M. & Robinson, C. V. The role of mass spectrometry in structure elucidation of dynamic protein complexes. *Annual Review of Biochemistry* **76,** 167–193 (2007).

36. Van Berkel, G. J., Zhou, F. & Aronson, J. T. Changes in bulk solution pH caused by the inherent controlled-current electrolytic process of an electrospray ion source. *International Journal of Mass Spectrometry and Ion Processes* **162,** 55–67. https://linkinghub.elsevier.com/retrieve/pii/S016811769604476X (1997).

37. Chernushevich, I. V., Bahr, U. & Karas, M. Nanospray 'taxation' and how to avoid it. *Rapid Communications in Mass Spectrometry* **18,** 2479–2485. https://onlinelibrary.wiley.com/doi/10.1002/rcm.1649 (2004).

38. Kaltashov, I. A. & Mohimen, A. Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. *Analytical Chemistry* **77,** 5370–5379. http://www.rcsb.org/pdb (2005).

39. Konermann, L. & Douglas, D. J. Unfolding of proteins monitored by electrospray ionization mass spectrometry: a comparison of positive and negative ion modes. *Journal of the American Society for Mass Spectrometry* **9,** 1248–1254. https://linkinghub.elsevier.com/retrieve/pii/S1525157812000967%20https://pubs.acs.org/doi/10.1016/S1044-0305%2898%2900103-2 (1998).

40. Winston, R. L. & Fitzgerald, M. C. Mass spectrometry as a readout of protein structure and function. *Mass Spectrometry Reviews* **16,** 165–179. https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1098-2787(1997)16:4%3C165::AID-MAS1%3E3.0.CO;2-F (1997).

41. Sobott, F., Hernández, H., McCammon, M. G., Tito, M. A. & Robinson, C. V. A Tandem Mass Spectrometer for Improved Transmission and Analysis of Large Macromolecular Assemblies. *Analytical Chemistry* **74,** 1402–1407. https://pubs.acs.org/doi/10.1021/ac0110552 (2002).

42. Light-Wahl, K. J., Schwartz, B. L. & Smith, R. D. Observation of the Noncovalent Quaternary Associations of Proteins by Electrospray Ionization Mass Spectrometry. *Journal of the American Chemical Society* **116,** 5271–5278. https://pubs.acs.org/sharingguidelines%20https://pubs.acs.org/doi/abs/10.1021/ja00091a035 (1994).

43. Winger, B. E., Light-Wahl, K. J., Ogorzalek Loo, R. R., Udseth, H. R. & Smith, R. D. Observation and implications of high mass-to-charge ratio ions from electrospray ionization mass spectrometry. *Journal of the American Society for Mass Spectrometry* **4,** 536–545. https://pubs.acs.org/doi/10.1016/1044-0305%2893%2985015-P (1993).

44. Krutchinsky, A. N., Chernushevich, I. V., Spicer, V. L., Ens, W. & Standing, K. G. Collisional damping interface for an electrospray ionization time-of-flight mass spectrometer. *Journal of the American Society for Mass Spectrometry* **9,** 569–579. https://pubs.acs.org/doi/10.1016/S1044-0305%2898%2900027-0 (1998).

45. Chernushevich, I. V. & Thomson, B. A. Collisional Cooling of Large Ions in Electrospray Mass Spectrometry. *Analytical Chemistry* **76,** 1754–1760. https://pubs.acs.org/doi/10.1021/ac035406j (2004).

46. Douglas, D. J. & French, J. B. Collisional focusing effects in radio frequency quadrupoles. *Journal of the American Society for Mass Spectrometry* **3,** 398–408. https://pubs.acs.org/doi/10.1016/1044-0305%2892%2987067-9 (1992).

47. Collings, B. A. & Douglas, D. J. An extended mass range quadrupole for electrospray mass spectrometry. *International journal of mass spectrometry and ion processes* **162,** 121–127 (1997).

48. Chung, E. W. *et al.* Probing the nature of interactions in SH2 binding interfaces—Evidence from electrospray ionization mass spectrometry. *Protein Science* **8,** 1962–1970 (1999).

49. Batey, J. H. The physics and technology of quadrupole mass spectrometers. *Vacuum* **101,** 410–415. https://linkinghub.elsevier.com/retrieve/pii/S0042207X13001541 (2014).

50. Paul, W. & Steinwedel, H. Ein neues massenspektrometer ohne magnetfeld. *Zeitschrift für Naturforschun A* **8,** 448–450 (1953).

51. Finnigan, R. E. Quadrupole mass spectrometers. *Analytical chemistry* **66,** 969A–975A (1994).

52. De Hoffmann, E. & Stroobant, V. *Mass spectrometry: principles and applications* (John Wiley & Sons, 2007).

53. Wolff, M. M. & Stephens, W. E. A Pulsed Mass Spectrometer with Time Dispersion. *Review of Scientific Instruments* **24,** 616–617. https://pubs.aip.org/rsi/article/24/8/616/441131/A-Pulsed-Mass-Spectrometer-with-Time-Dispersion (1953).

54. Boesl, U. Time-of-flight mass spectrometry: Introduction to the basics. *Mass Spectrometry Reviews* **36,** 86–109. https://onlinelibrary.wiley.com/doi/10.1002/mas.21520 (2017).

55. Wiley, W. C. & McLaren, I. H. Time-of-Flight Mass Spectrometer with Improved Resolution. *Review of Scientific Instruments* **26,** 1150–1157. https://pubs.aip.org/rsi/article/26/12/1150/298880/Time-of-Flight-Mass-Spectrometer-with-Improved (1955).

56. Mamyrin, B. A., Karataev, V. I., Shmikk, D. V. & Zagulin, V. A. The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Sov Phys JETP* **37,** 45–48. http://www.google.com/search?client=safari&rls=en-us&q=The+mass-reflectron,+a+new+nonmagnetic+time-of-flight+mass+spectrometer+with+high+resolution&ie=UTF-8&oe=UTF-8 (1973).

57. Guilhaus, M., Selby, D. & Mlynski, V. Orthogonal acceleration time-of-flight mass spectrometry. *Mass Spectrometry Reviews* **19,** 65–107. https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1098-2787(2000)19:2%3C65::AID-MAS1%3E3.0.CO;2-E (2000).

58. Chernushevichfn, I. V., Verentchikovfn, A. N., Ens, W. & Standingcor, K. G. Effect of ion-molecule collisions in the vacuum chamber of an electrospray time-of-flight mass spectrometer on mass spectra of proteins. *Journal of the American Society for Mass Spectrometry* **7,** 342–349. https://pubs.acs.org/sharingguidelines%20https://pubs.acs.org/doi/10.1016/1044-0305%2895%2900681-8 (1996).

59. Makarov, A. *et al.* Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *Analytical Chemistry* **78,** 2113–2120. https://pubs.acs.org/doi/10.1021/ac0518811 (2006).

60. Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Analytical Chemistry* **72,** 1156–1162. https://pubs.acs.org/doi/10.1021/ac991131p (2000).

61. Makarov, A. A. *Mass spectrometer* 1999.

62. Makarov, A., Hardman, M. E., Schwartz, J. C. & Senko, M. W. *Mass spectrometry method and apparatus* 2005.

63. Fort, K. L. *et al.* Expanding the structural analysis capabilities on an Orbitrap-based mass spectrometer for large macromolecular complexes. *The Analyst* **143,** 100–105. http://xlink.rsc.org/?DOI=C7AN01629H (2018).

64. Benesch, J. L., Aquilina, J. A., Ruotolo, B. T., Sobott, F. & Robinson, C. V. Tandem Mass Spectrometry Reveals the Quaternary Organization of Macromolecular Assemblies. *Chemistry & Biology* **13,** 597–605. https://linkinghub.elsevier.com/retrieve/pii/S1074552106001463 (2006).

65. Jennings, K. R. The changing impact of the collision-induced decomposition of ions on mass spectrometry. *International Journal of Mass Spectrometry* **200,** 479–493. https://linkinghub.elsevier.com/retrieve/pii/S1387380600003250 (2000).

66. Barth, M. & Schmidt, C. Native mass spectrometry—A valuable tool in structural biology. *Journal of Mass Spectrometry* **55.** https://onlinelibrary.wiley.com/doi/10.1002/jms.4578 (2020).

67. Olsen, J. V. *et al.* Higher-energy C-trap dissociation for peptide modification analysis. *Nature Methods* **4,** 709–712. https://www.nature.com/articles/nmeth1060 (2007).

68. Vinciauskaite, V. & Masson, G. R. Fundamentals of HDX-MS. *Essays in Biochemistry* **67** (eds Britt, H., Beveridge, R. & Calabrese, A.) 301–314. https://portlandpress.com/essaysbiochem/article/67/2/301/231942/Fundamentals-of-HDX-MS (2023).

69. Wales, T. E. & Engen, J. R. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrometry Reviews* **25,** 158–170. https://onlinelibrary.wiley.com/doi/10.1002/mas.20064 (2006).

70. Dyson, H. J. & Wright, P. E. Unfolded Proteins and Protein Folding Studied by NMR. *Chemical Reviews* **104,** 3607–3622. https://pubs.acs.org/doi/10.1021/cr030403s (2004).

71. Green, M. K. & Lebrilla, C. B. Ion-molecule reactions as probes of gas-phase structures of peptides and proteins. *Mass Spectrometry Reviews* **16,** 53–71. https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1098-2787(1997)16:2%3C53::AID-MAS1%3E3.0.CO;2-8 (1997).

72. Hamuro, Y. & Coales, S. J. Optimization of Feasibility Stage for Hydrogen/Deuterium Exchange Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **29,** 623–629. https://pubs.acs.org/doi/10.1007/s13361-017-1860-3 (2018).

73. Morgan, C. R. & Engen, J. R. Investigating Solution-Phase Protein Structure and Dynamics by Hydrogen Exchange Mass Spectrometry. *Current Protocols in Protein Science* **58.** https://onlinelibrary.wiley.com/doi/10.1002/0471140864.ps1706s58 (2009).

74. Burns, K. M., Rey, M., Baker, C. A. & Schriemer, D. C. Platform Dependencies in Bottom-up Hydrogen/Deuterium Exchange Mass Spectrometry. *Molecular & Cellular Proteomics* **12,** 539–548. http://dx.doi.org/10.1074/mcp.M112.023770%20https://linkinghub.elsevier.com/retrieve/pii/S1535947620332618 (2013).

75. Mistarz, U. H. & Rand, K. D. Installation, validation, and application examples of two instrumental setups for gas-phase HDX-MS analysis of peptides and proteins. *Methods* **144,** 113–124. https://linkinghub.elsevier.com/retrieve/pii/S1046202317304620 (2018).

76. Breuker, K. & McLafferty, F. W. Stepwise evolution of protein native structure with electrospray into the gas phase, 10- 12 to 102 s. *Proceedings of the National Academy of Sciences* **105,** 18145–18152. www.pnas.orgcgidoi10.1073pnas.0807005105%20https://pnas.org/doi/full/10.1073/pnas.0807005105 (2008).

77. Masson, G. R. *et al.* Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nature Methods* **16,** 595–602. https://www.nature.com/articles/s41592-019-0459-y (2019).

78. Perkins, D. N., Pappin, D. J. C., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20,** 3551–3567. https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1522-2683(19991201)20:18%3C3551::AID-ELPS3551%3E3.0.CO;2-2 (1999).

79. Claesen, J. & Burzykowski, T. Computational methods and challenges in hydrogen/deuterium exchange mass spectrometry. *Mass Spectrometry Reviews* **36,** 649–667. https://onlinelibrary.wiley.com/doi/10.1002/mas.21519 (2017).

80. Hvidt, A. & Linderstrøm-Lang, K. The kinetics of the deuterium exchange of insulin with D2O. An amendment. *Biochimica et Biophysica Acta* **16,** 168–169. https://linkinghub.elsevier.com/retrieve/pii/0006300255902006 (1955).

81. Berger, A. & LinderstrØm-Lang, K. Deuterium exchange of poly-DL-alanine in aqueous solution. *Archives of Biochemistry and Biophysics* **69,** 106–118. https://linkinghub.elsevier.com/retrieve/pii/S0003986122001242 (1957).

82. Pan, J., Han, J., Borchers, C. H. & Konermann, L. Hydrogen/Deuterium Exchange Mass Spectrometry with Top-Down Electron Capture Dissociation for Characterizing Structural Transitions of a 17 kDa Protein. *Journal of the American Chemical Society* **131,** 12801–12808. https://pubs.acs.org/doi/10.1021/ja904379w (2009).

83. Rand, K. D., Zehl, M. & Jørgensen, T. J. D. Measuring the Hydrogen/Deuterium Exchange of Proteins at High Spatial Resolution by Mass Spectrometry: Overcoming Gas-Phase Hydrogen/Deuterium Scrambling. *Accounts of Chemical Research* **47,** 3018–3027. https://pubs.acs.org/doi/10.1021/ar500194w (2014).

84. Konermann, L., Pan, J. & Liu, Y.-H. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* **40,** 1224–1234. http://xlink.rsc.org/?DOI=C0CS00113A (2011).

85. Greisch, J.-F. *et al.* Expanding the mass range for UVPD-based native top-down mass spectrometry. *Chemical Science* **10,** 7163–7171. http://xlink.rsc.org/?DOI=C9SC01857C (2019).

86. Brodbelt, J. S., Morrison, L. J. & Santos, I. Ultraviolet Photodissociation Mass Spectrometry for Analysis of Biological Molecules. *Chemical Reviews* **120,** 3328–3380. https://pubs.acs.org/doi/10.1021/acs.chemrev.9b00440 (2020).

87. Brodie, N. I. *et al.* Top-Down Hydrogen–Deuterium Exchange Analysis of Protein Structures Using Ultraviolet Photodissociation. *Analytical Chemistry* **90,** 3079–3082. https://pubs.acs.org/doi/10.1021/acs.analchem.7b03655 (2018).

88. Ravenhill, B. J. *et al.* The Cargo Receptor NDP52 Initiates Selective Autophagy by Recruiting the ULK Complex to Cytosol-Invading Bacteria. *Molecular Cell* **74,** 320–329.e6. https://doi.org/10.1016/j.molcel.2019.01.041%20https://linkinghub.elsevier.com/retrieve/pii/S1097276519300619 (2019).

89. Ohashi, Y. *et al.* Membrane characteristics tune activities of endosomal and autophagic human VPS34 complexes. *eLife* **9,** 1–28. https://elifesciences.org/articles/58281 (2020).

90. Rob, T. *et al.* Measuring Dynamics in Weakly Structured Regions of Proteins Using Microfluidics-Enabled Subsecond H/D Exchange Mass Spectrometry. *Analytical Chemistry* **84,** 3771–3779. https://pubs.acs.org/doi/10.1021/ac300365u (2012).

91. Seetaloo, N., Zacharopoulou, M., Stephens, A. D., Kaminski Schierle, G. S. & Phillips, J. J. Millisecond Hydrogen/Deuterium-Exchange Mass Spectrometry Approach to Correlate Local Structure and Aggregation in α-Synuclein. *Analytical Chemistry* **94,** 16711–16719. https://pubs.acs.org/doi/10.1021/acs.analchem.2c03183 (2022).

92. Kish, M. *et al.* Online Fully Automated System for Hydrogen/Deuterium-Exchange Mass Spectrometry with Millisecond Time Resolution. *Analytical Chemistry* **95,** 5000–5008. https://pubs.acs.org/doi/10.1021/acs.analchem.2c05310 (2023).

93. Chalmers, M. J. *et al.* Probing protein ligand interactions by automated hydrogen/deuterium exchange mass spectrometry. *Analytical Chemistry* **78,** 1005–1014 (2006).

94. Rapid Hydrogen–Deuterium Exchange in Liquid Droplets. *Journal of the American Chemical Society* **139,** 6851–6854. https://pubs.acs.org/doi/10.1021/jacs.7b03541 (2017).

95. James, E. I., Murphree, T. A., Vorauer, C., Engen, J. R. & Guttman, M. Advances in Hydrogen/Deuterium Exchange Mass Spectrometry and the Pursuit of Challenging Biological Systems. *Chemical Reviews* **122,** 7562–7623. https://pubs.acs.org/doi/10.1021/acs.chemrev.1c00279 (2022).

96. Konermann, L. & Simmons, D. A. Protein-folding kinetics and mechanisms studied by pulse-labeling and mass spectrometry. *Mass Spectrometry Reviews* **22,** 1–26. https://onlinelibrary.wiley.com/doi/10.1002/mas.10044 (2003).

97. Molday, R., Englander, S. & Kallen, R. Primary structure effects on peptide group hydrogen exchange. *Biochemistry* **11,** 150–158. https://pubs.acs.org/sharingguidelines%20https://onlinelibrary.wiley.com/doi/10.1002/prot.340170110 (1972).

98. Nguyen, D., Mayne, L., Phillips, M. C. & Walter Englander, S. Reference Parameters for Protein Hydrogen Exchange Rates. *Journal of the American Society for Mass Spectrometry* **29,** 1936–1939 (2018).

99. Oganesyan, I., Lento, C. & Wilson, D. J. Contemporary hydrogen deuterium exchange mass spectrometry. *Methods* **144,** 27–42. https://linkinghub.elsevier.com/retrieve/pii/S1046202317304528 (2018).

100. Miller, D. W. & Dill, K. A. A statistical mechanical model for hydrogen exchange in globular proteins. *Protein Science* **4,** 1860–1873. https://onlinelibrary.wiley.com/doi/10.1002/pro.5560040921 (1995).

101. Xiao, H. Mapping protein energy landscapes with amide hydrogen exchange and mass spectrometry: I. A generalized model for a two-state protein and comparison with experiment. *Protein Science* **14,** 543–557. http://doi.wiley.com/10.1110/ps.041001705 (2005).

102. Guttman, M., Weis, D. D., Engen, J. R. & Lee, K. K. Analysis of Overlapped and Noisy Hydrogen/Deuterium Exchange Mass Spectra. *Journal of the American Society for Mass Spectrometry* **24,** 1906–1912. https://pubs.acs.org/doi/10.1007/s13361-013-0727-5 (2013).

103. Wu, D. & Piszczek, G. Standard protocol for mass photometry experiments. *European Biophysics Journal* **50,** 403–409. https://link.springer.com/10.1007/s00249-021-01513-9 (2021).

104. Asor, R. & Kukura, P. Characterising biomolecular interactions and dynamics with mass photometry. *Current Opinion in Chemical Biology* **68,** 102132. https://doi.org/10.1016/j.cbpa.2022.102132%20https://linkinghub.elsevier.com/retrieve/pii/S1367593122000175 (2022).

105. Quantitative mass imaging of single biological macromolecules. *Science* **360,** 423–427. https://www.science.org/doi/10.1126/science.aar5839 (2018).

106. Young, G. & Kukura, P. Interferometric Scattering Microscopy. *Annual Review of Physical Chemistry* **70,** 301–322. https://www.annualreviews.org/doi/10.1146/annurev-physchem-050317-021247 (2019).

107. Ortega-Arroyo, J. & Kukura, P. Interferometric scattering microscopy (iSCAT): new frontiers in ultrafast and ultrasensitive optical microscopy. *Physical Chemistry Chemical Physics* **14,** 15625. http://xlink.rsc.org/?DOI=c2cp41013c (2012).

108. Eaton, G. R. & Eaton, S. S. *Foundations of modern EPR* (World Scientific, 1998).

109. Rhodes, C. J. Applications of EPR in the Environmental Sciences. *EPR of Free Radicals in Solids II: Trends in Methods and Applications,* 279–310 (2012).

110. Spasojević, I. Free radicals and antioxidants at a glance using EPR spectroscopy. *Critical reviews in clinical laboratory sciences* **48,** 114–142 (2011).

111. Swartz, H. M. *et al.* Clinical EPR. *Academic Radiology* **21,** 197–206. https://linkinghub.elsevier.com/retrieve/pii/S1076633213004996 (2014).

112. Jeschke, G. The contribution of modern EPR to structural biology. *Emerging Topics in Life Sciences* **2** (eds Ubbink, M. & Perrakis, A.) 9–18. https://portlandpress.com/emergtoplifesci/article/2/1/9/76649/The-contribution-of-modern-EPR-to-structural (2018).

113. Chechik, V., Carter, E. & Murphy, D. *Electron Paramagentic Resonance* 1st editio (Oxford University Press, Oxford, 2015).

114. Stoll, S. & Goldfarb, D. *EPR spectroscopy: fundamentals and methods* (John Wiley & Sons, 2018).

115. Eaton, G. R., Eaton, S. S., Barr, D. P. & Weber, R. T. *Quantitative EPR* (Springer Science & Business Media, 2010).

116. Schweiger, A. & Jeschke, G. *Principles of pulse electron paramagnetic resonance* (Oxford University Press, 2001).

117. Schiemann, O. *et al.* Benchmark Test and Guidelines for DEER/PELDOR Experiments on Nitroxide-Labeled Biomolecules. *Journal of the American Chemical Society* **143,** 17875–17890. https://pubs.acs.org/doi/10.1021/jacs.1c07371 (2021).

118. Jeschke, G. & Polyhach, Y. Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Physical Chemistry Chemical Physics* **9,** 1895. http://xlink.rsc.org/?DOI=b614920k (2007).

119. Jeschke, G. Distance Measurements in the Nanometer Range by Pulse EPR. *ChemPhysChem* **3,** 927–932. https://onlinelibrary.wiley.com/doi/10.1002/1439-7641(20021115)3:11%3C927::AID-CPHC927%3E3.0.CO;2-Q (2002).

120. Hollingsworth, S. A. & Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **99,** 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011%20https://linkinghub.elsevier.com/retrieve/pii/S0896627318306846 (2018).

121. Alder, B. J. & Wainwright, T. E. Phase Transition for a Hard Sphere System. *The Journal of Chemical Physics* **27,** 1208–1209. https://pubs.aip.org/jcp/article/27/5/1208/204728/Phase-Transition-for-a-Hard-Sphere-System (1957).

122. McCammon, J. A., Gelin, B. R. & Karplus, M. Dynamics of folded proteins. *Nature* **267,** 585–590. https://www.nature.com/articles/267585a0.pdf%0Ahttp://www.msg.ucsf.edu/BI204/Discussion/Q1/Jacobson/MacCammon_1977.pdf%20https://www.nature.com/articles/267585a0 (1977).

123. Karplus, M. & McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nature Structural Biology* **9,** 646–652. https://www.jstage.jst.go.jp/article/jsmemag/116/1131/116_KJ00008530809/_article/-char/ja/%20https://www.nature.com/doifinder/10.1038/nsb0902-646 (2002).

124. Warshel, A. & Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology* **103,** 227–249. https://linkinghub.elsevier.com/retrieve/pii/0022283676903119 (1976).

125. Senn, H. M. & Thiel, W. QM/MM Methods for Biomolecular Systems. *Angewandte Chemie International Edition* **48,** 1198–1229. https://onlinelibrary.wiley.com/doi/10.1002/anie.200802019 (2009).

126. Jorgensen, W. L. & Tirado-Rives, J. Monte Carlo vs Molecular Dynamics for Conformational Sampling. *The Journal of Physical Chemistry* **100,** 14508–14513. https://pubs.acs.org/doi/10.1021/jp960880x (1996).

127. Jensen, F. *Introduction to computational chemistry* 2nd Edition (John Wiley & Sons, 2017).

128. Guvench, O. & MacKerell, A. D. in *Methods in Molecular Biology: Comparison of Protein Force Fields for Molecular Dynamics Simulations* 63–88 (2008).

129. Monticelli, L. & Tieleman, D. P. in *Biomolecular simulations: Methods and protocols* 197–213 (Springer, 2013).

130. Cornell, W. D. *et al.* A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* **117,** 5179–5197. https://pubs.acs.org/doi/abs/10.1021/ja00124a002 (1995).

131. MacKerell, A. D. *et al.* All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *The Journal of Physical Chemistry B* **102,** 3586–3616. https://pubs.acs.org/doi/10.1021/jp973084f (1998).

132. Oostenbrink, C., Villa, A., Mark, A. E. & Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry* **25,** 1656–1676. https://onlinelibrary.wiley.com/doi/10.1002/jcc.20090 (2004).

133. Cramer, C. J. *Essentials of computational chemistry: theories and models* 2nd Edition (John Wiley & Sons, 2013).

134. Jorgensen, W. L., Tirado-rives, J. & Fields, F. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proceedings of the National Academy of Sciences* **102,** 6665–6670 (2005).

135. Barclay, P. L. & Zhang, D. Z. Periodic boundary conditions for arbitrary deformations in molecular dynamics simulations. *Journal of Computational Physics* **435,** 110238. https://doi.org/10.1016/j.jcp.2021.110238 (2021).

136. Hansson, T., Oostenbrink, C. & Gunsteren, W. F. V. Molecular dynamics simulations. *Current Opinion in Structural Biology* **12,** 190–196 (2002).

137. Schreiber, H. & Steinhauser, O. Cutoff size does strongly influence molecular dynamics results on solvated polypeptides. *Biochemistry* **31,** 5856–5860 (1992).

138. Essmann, U. *et al.* A smooth particle mesh Ewald method. *The Journal of chemical physics* **103,** 8577–8593 (1995).

139. Ke, Q., Gong, X., Liao, S., Duan, C. & Li, L. Effects of thermostats/barostats on physical properties of liquids by molecular dynamics simulations. *Journal of Molecular Liquids* **365,** 120116. https://doi.org/10.1016/j.molliq.2022.120116%20https://linkinghub.elsevier.com/retrieve/pii/S0167732222016555 (2022).

All the flowers that you planted, mama, in the backyard
All died when you went away
I know that living with you, baby, was sometimes hard
But I'm willing to give it another try

Nothing compares
Nothing compares to you

— from Sinéad O'Connor's cover of *Nothing Compares 2 U*

3

Conformational changes through the lens of HDX-MS

Contents

3.1 Introduction 87
3.2 Light-induced conformational changes in European robin cryptochrome 89
3.3 Introducing the *Er*Cry4a W369F mutant 93
3.3.1 Could different buffer conditions improve our results? . 101
3.3.2 Tracking down the C-terminal tail 104
3.3.3 Concluding discussion 107
3.3.4 Final remarks 109
3.4 Material and Methods 110
3.4.1 Material 110
3.4.2 HDX-MS 111
3.4.3 Native MS 113
3.4.4 Data analysis 113
References 114

3.1 Introduction

Conformational changes in a protein can be crucial to its function [1] and elucidating them provides information on which parts of the protein might be responsive

to external changes in environmental conditions and to what extent these parts change. In this chapter the changes in solvent accessibility which may be the result of conformational changes in European robin cryptochrome 4a and one of its functionally interesting mutants are explored and analysed with respect to their potential biological function.

This chapter also explores experimental difficulties and leads to a set of possible improvements which were subsequently implemented and led to the results presented in chapter 4. Therefore, only the work presented in the latter is currently being prepared for publication whereas the work presented in this chapter will not be published although it has been instrumental in paving the way towards acquiring data of improved quality.

European robin or *Erithacus rubecula* cryptochrome 4a (*Er*Cry4a) is of investigative interest as data acquired from this protein could be directly compared with or related to a wealth of published data on these night-migratory birds. Several research groups conducted behavioural experiments on European robins over the years [2–6] and more recently, purified *Er*Cry4a has been studied experimentally [7–9]. We have received all *Er*Cry4a samples from our collaborators in the group of Henrik Mouritsen at the University in Oldenburg and submitted them to different light and magnetic field conditions, analogous to what had been done in earlier spectroscopic studies [7]. However, the aim of this work was to elucidate the downstream conformational changes that the formation of radical pairs within the protein might ultimately cause.

At the time the present work started, there were only a few publications describing how HDX-MS had been used to analyse cryptochromes, one of them looking at the bifunctional cryptochrome/photolyase animal-like cryptochrome from the green algae *Chlamydomonas reinhardtii* (*Cra*CRY) [10]. Similarly to the already ongoing work in our research group on *Drosophila melanogaster* cryptochrome (*Dm*Cry) [11] (for more details see chapter 6), the authors of this study also used a light chamber on an HDX-MS setup to compare the light and dark states of the protein with each other and found that the C-terminal α-helix α22 loses

conformational restraint upon illumination with light [10]. In 2020, the same research group found a novel subgroup of the photolyase/cryptochrome superfamily called NewPHLs and also used HDX-MS to investigate its specific characteristics [12]. Some work using HDX-MS has also been performed on *Arabidopsis thaliana* cryptochrome 2, a protein involved in the regulation of plant development, analysing conformational changes induced by the binding of adenosine triphosphate (ATP) [13].

However, at the time of writing no similar investigation had been performed on any avian cryptochrome species leading us to believe that an investigation of *Er*Cry4a might bring to light so far unknown information about the mechanism of magnetoreception.

3.2 Light-induced conformational changes in European robin cryptochrome

The first attempt to work with the wild type (WT) version of *Er*Cry4a for the present work yielded very unexpected results identifying a problem with the way the experiments were performed. This problem was subsequently fixed. The results of this problematic experiment are displayed as the relative fractional uptake across the entire protein sequence in Figure 3.1A and as difference uptake plots in Figure 3.1B with the sequence coverage as depicted in the Appendix Figure A.1. Four different states were investigated, all at concentrations in the range 10-16 µM, the light state (LS) where proteins were subjected to at least 20 min of exposure to blue light and the dark state (DS) where the proteins were kept for 20 min in complete darkness, as well as both of these states with the addition of a permanent magnet applying an additional, external magnetic field in the mT range, denoted LSM and DSM, respectively. For these states, a permanent magnet with a strength of 160 mT was chosen, leading to a gradient across the sample with values ranging from approximately 18-25 mT. These values are stronger than the Earth's magnetic field and are intended to reach the point at which the magnetic field effect saturates. This is thought to increase the chances of measuring a discernible difference from non-magnet states in an experimental setup which

measures the surface accessibility under equilibrium conditions. The uptake plots indicate similar uptake in similar regions of all four investigated states with the C-terminal region consistently displaying large uptake values and a saturation that is already reached at the earliest time points measured. The uptake difference plots show the largest differences when comparing either of the two illumination states with their respective magnet state. This is surprising as from previous literature [10, 14, 15] it would have been expected to see notable differences when comparing the LS with the DS (LS-DS), if not much larger than those seen with the addition of a magnet (LSM-LS and DSM-DS).

Figure 3.2 displays peptides that stood out during the analysis of the data displayed in Figure 3.1. Again, the main differences seen are between a given state and its magnet-present counterpart, as expected from the results presented in Figure 3.1. Due to the inconsistent nature of the HDX-MS setup's workflow due to frequent and often unpredictable breakdowns of various instrumental components, it was necessary to perform this experiment as time-efficiently as possible. The most efficient practice at the time was to perform the experimental procedures including all repeats for each state separately one after another. They were performed in the following order: DS, LS, LSM, and DSM. Notably, the first four peptides displayed in Figure 3.2 show that the two states investigated last are the ones with the highest uptakes. The reason for this is suspected to be an instability of the protein over time while being kept at room temperature for several hours. Previous experiments performed on this or similar proteins (*Arabidopsis thaliana*, chicken, and pigeon cryptochromes) did not indicate the same instability [7]. However, these experiments were mostly performed at much lower temperatures which could account for the differences in stability. As a change in species and/or a change in buffer conditions and temperature could lead to a different protein stability, the results pointed out the need to improve the experimental workflow by changing the order in which the experimental repeats for the different states were recorded and by reducing the time the proteins spend at room temperature.

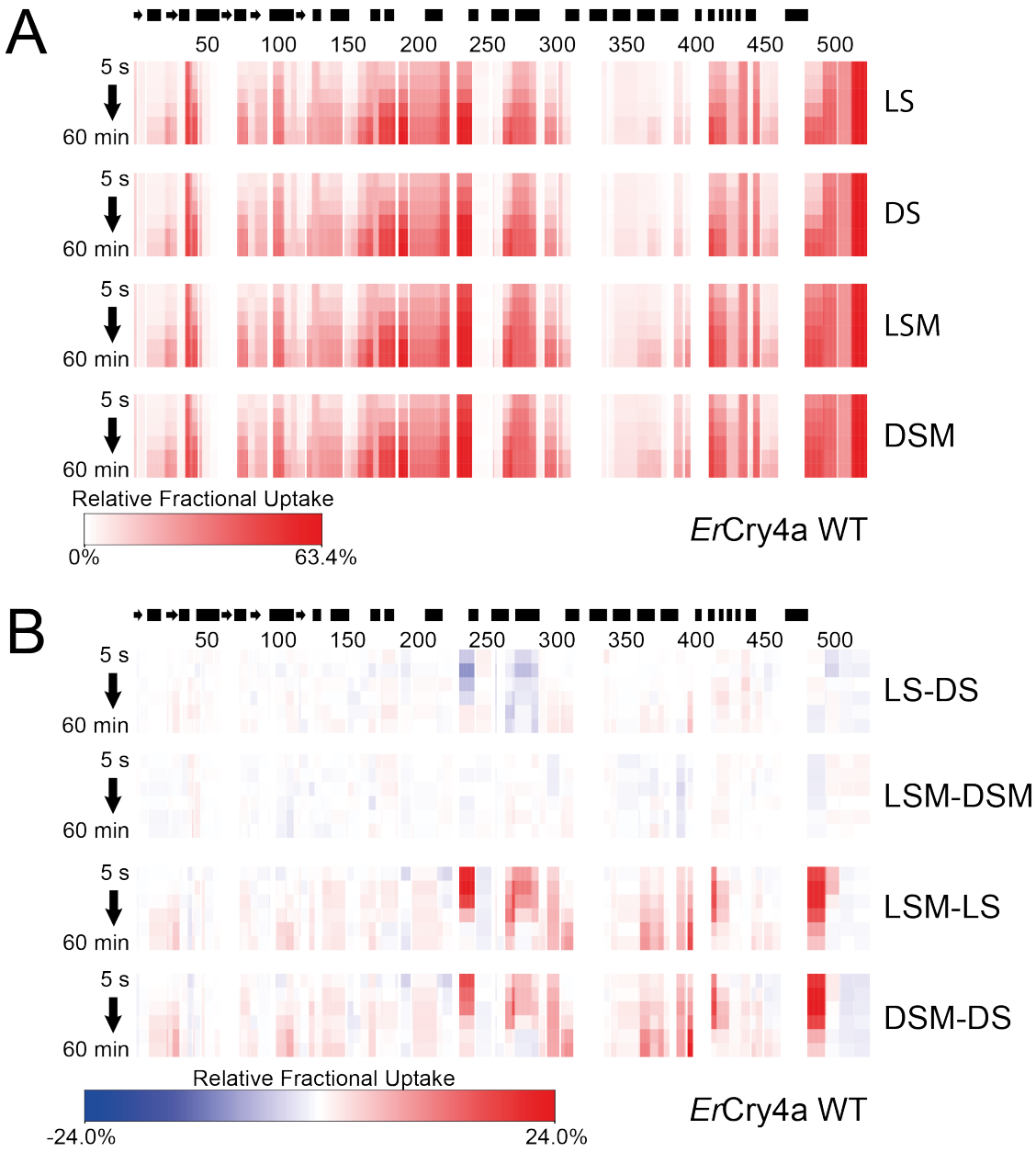

Figure 3.1: Relative fractional uptake of *Er*Cry4a WT. (**A**) Relative fractional uptake across the entire sequence measured at the peptide level. (**B**) Relative fractional uptake differences between states as indicated. As per usual convention, the N-terminus is situated on the left side and the C-terminus on the right. The time points displayed for both (A) and (B) are 5 s, 15 s, 1 min, 5 min, 30 min, and 1 hr. The data are based on a total of 112 peptides, a sequence coverage of 85.6 %, and an average redundancy of 2.44. The relative fractional uptake is calculated by dividing the deuterium level in Da by the total number of backbone amide hydrogens that could have become deuterated (which is equal to the number of amino acids, with the exception of proline residues and N-terminal residues) [16, 17]. For overlapping peptides, the data displayed is determined by the shortest peptide. Beta sheets (arrows) and alpha-helices (rectangles) are indicated above the sequences [7]. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet. Unexpectedly, the LSM-LS and DSM-DS differences are larger than the LS-DS and LSM-DSM differences. The reason for this is suspected to be protein instability. For more details, see main text.

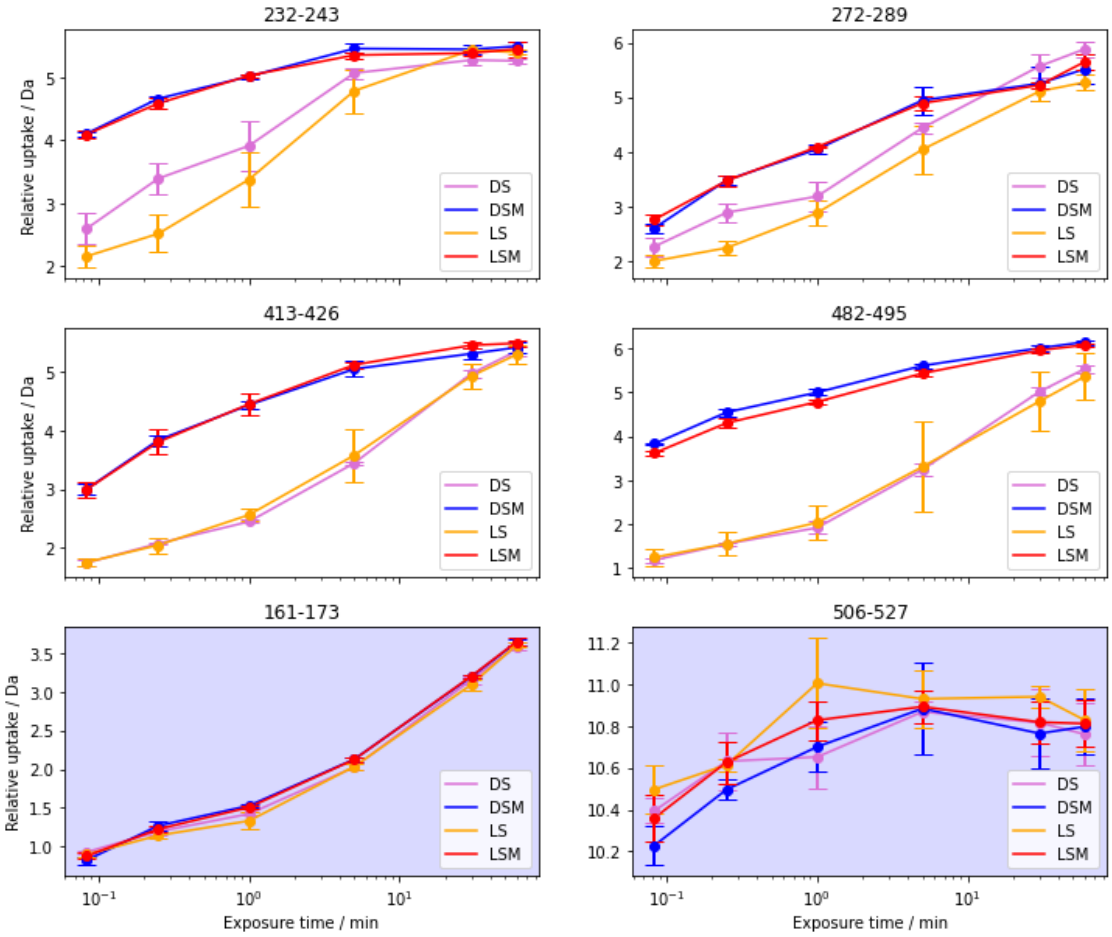

Figure 3.2: Relative uptake of representative *Er*Cry4a WT peptides. Peptide uptake curves and standard deviations for deuterium incorporation of six representative peptides. The first four relative uptake plots are representative of regions in which large changes between states were observed. The relative uptake plots highlighted in blue are examples of peptides with no or less clear uptake differences between states. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

Figure 3.2 also includes the peptides 161-173 and 506-527, highlighted in blue. Peptide 161-173 is an example of a protein region that experiences similar uptake over time for all four states. The peptides covering these regions generally seem to have a lower maximum uptake than those which experience any differences between states and are therefore less solvent-accessible. Peptide 506-527 is a representative peptide of those covering the C-terminal region. It shows a high overall uptake for all four states indicating a flexible and/or rather solvent exposed area.

3.3 Introducing the *Er*Cry4a W369F mutant

The *Er*Cry4a W369F mutant has a tryptophan residue replaced with a less redox-active phenylalanine as done previously [7, 9, 18], to block the electron transfer pathway from the 4th Trp member of the Trp tetrad and shorten the longest inter-radical distance to the distance between the flavin cofactor and the 3rd Trp member. As it was suggested that the 3rd and 4th members of the Trp tetrad could have different functions with the first being be more important for sensing while the latter could be more important for signalling [7, 19], the investigation of differences in conformational changes of this particular mutant were of interest.

The experiment for this mutant was performed during the same two weeks as the experiment on *Er*Cry4a WT discussed above. Therefore, the same experimental shortcomings and similar results after data analysis were expected. However, the trend of pronounced differences in uptake scaling with the order of acquisition were not seen for this mutant as displayed in Figure 3.5. This points towards its greater stability and structural integrity in comparison to the WT.

Figure 3.3 shows the relative fractional uptake across the entire sequence in (A) and the relative fractional uptake differences between the different investigated states in (B) (the corresponding peptide coverage can be found in Appendix Figure A.1). The former shows that overall the same regions experience similar levels of uptake irrespective of the state that was investigated, similar to the corresponding results found for *Er*Cry4a WT. Part (B) of the same figure shows large differences for some regions of the sequence when considering LS-DS differences with the LSM-LS and DSM-DS differences being rather marginal and less consistent over time per protein section. This is the opposite of what had been found for *Er*Cry4a WT. Some of the uptake and uptake difference values have been displayed on a 3D AlphaFold [20] structure of *Er*Cry4a WT in Figure 3.4 for easier visualisation of the results.

These results have to be interpreted with caution in the light of the results obtained for the WT protein and need verification by performing biological repeats. However, it is interesting to note that when focusing on the regions displaying the greatest differences in Figure 3.3B, one can see almost exclusively regions that

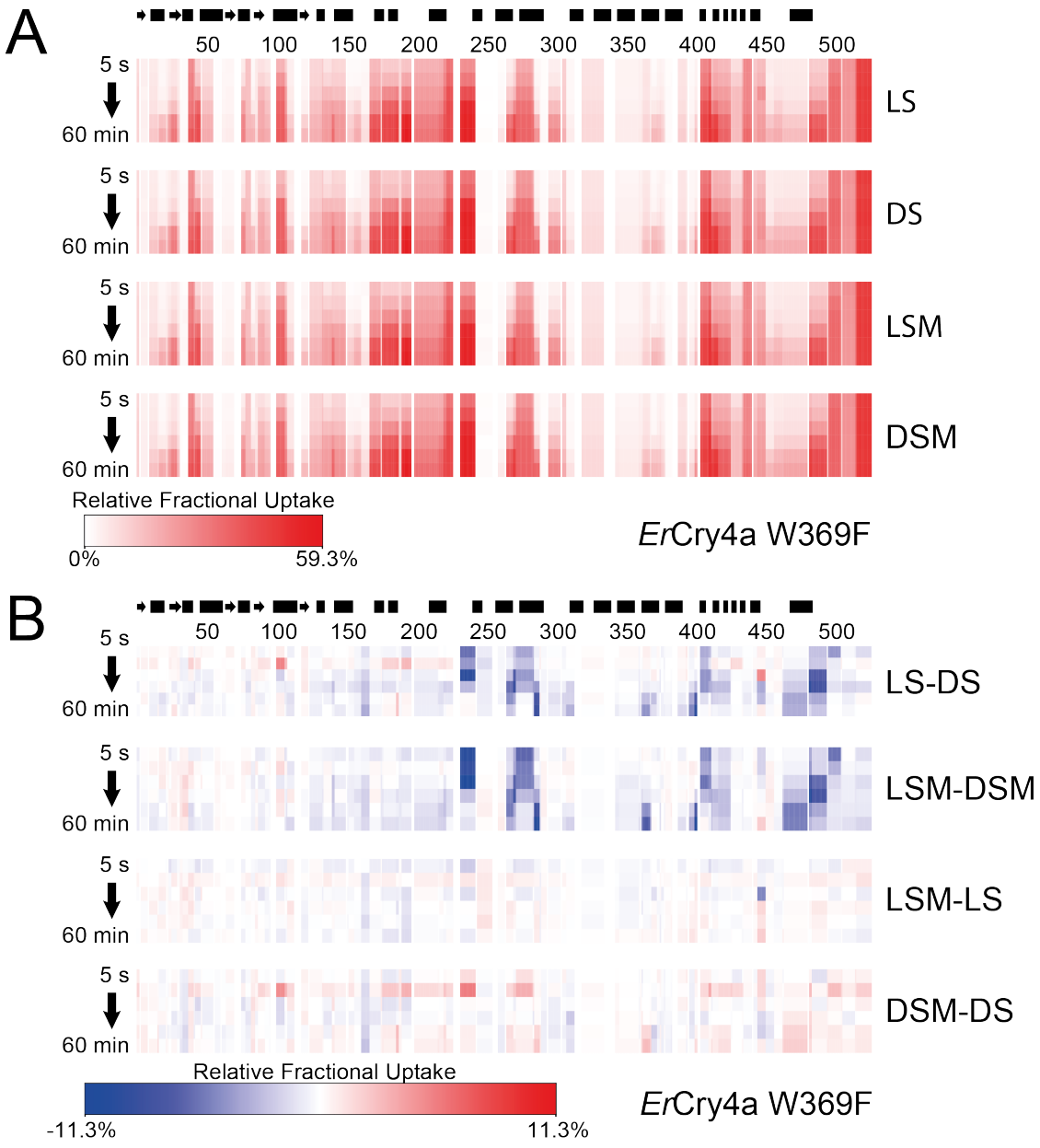

Figure 3.3: Relative fractional uptake of *Er*Cry4a W369F. (**A**) Relative fractional uptake across the entire sequence measured at the peptide level. (**B**) Relative fractional uptake differences between states as indicated. The time points displayed for both (A) and (B) are 5 s, 15 s, 1 min, 5 min, 30 min, and 1 hr. The data are based on a total of 138 peptides, a sequence coverage of 93.6 %, and an average redundancy of 2.83. All additional information can be found in the caption of Figure 3.1.

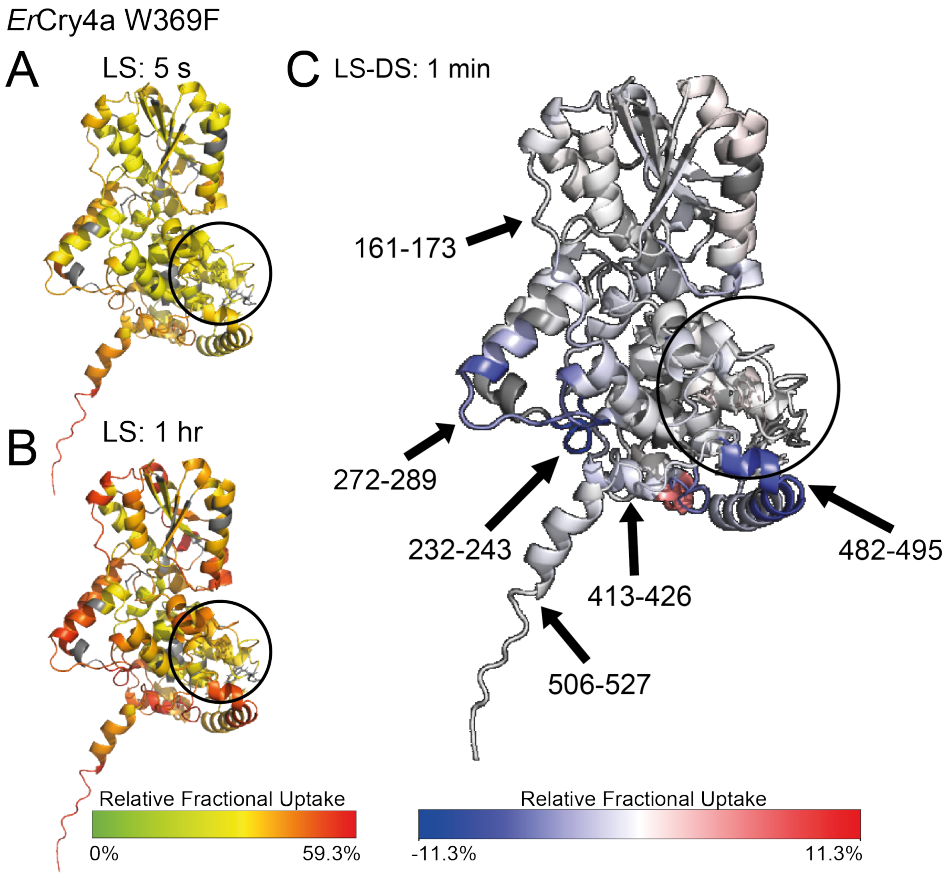

Figure 3.4: Relative fractional uptake of *Er*Cry4a W369F displayed on AlphaFold structures. (**A**) Relative fractional uptake across the entire protein sequence after 5 s. (**B**) Relative fractional uptake across the entire protein sequence after 1 hr of labelling. (**C**) Relative fractional uptake differences between LS and DS after 1 min of labelling. Structures were generated by Navoneel Sen using AlphaFold [20]. The area of the tryptophan tetrad is encircled in black. DS - dark state, LS - light state.

experience protection (blue) from surface accessibility upon illumination with blue light. This is also visible in Figure 3.5 which focuses on the peptides 232-243, 272-289, 413-426, and 482-495 as representative examples. No neighbouring regions experience de-protection which one might expect for larger conformational changes involving protection events. This point is further discussed in chapter 4. The same peptides as displayed in Figure 3.2 were chosen for a direct comparison of the differences between states. Therefore, Figure 3.5 also features two further peptides, 161-173 and 506-527, with the former being a representative example of a protein region where all states experience a similar uptake over time and the latter an

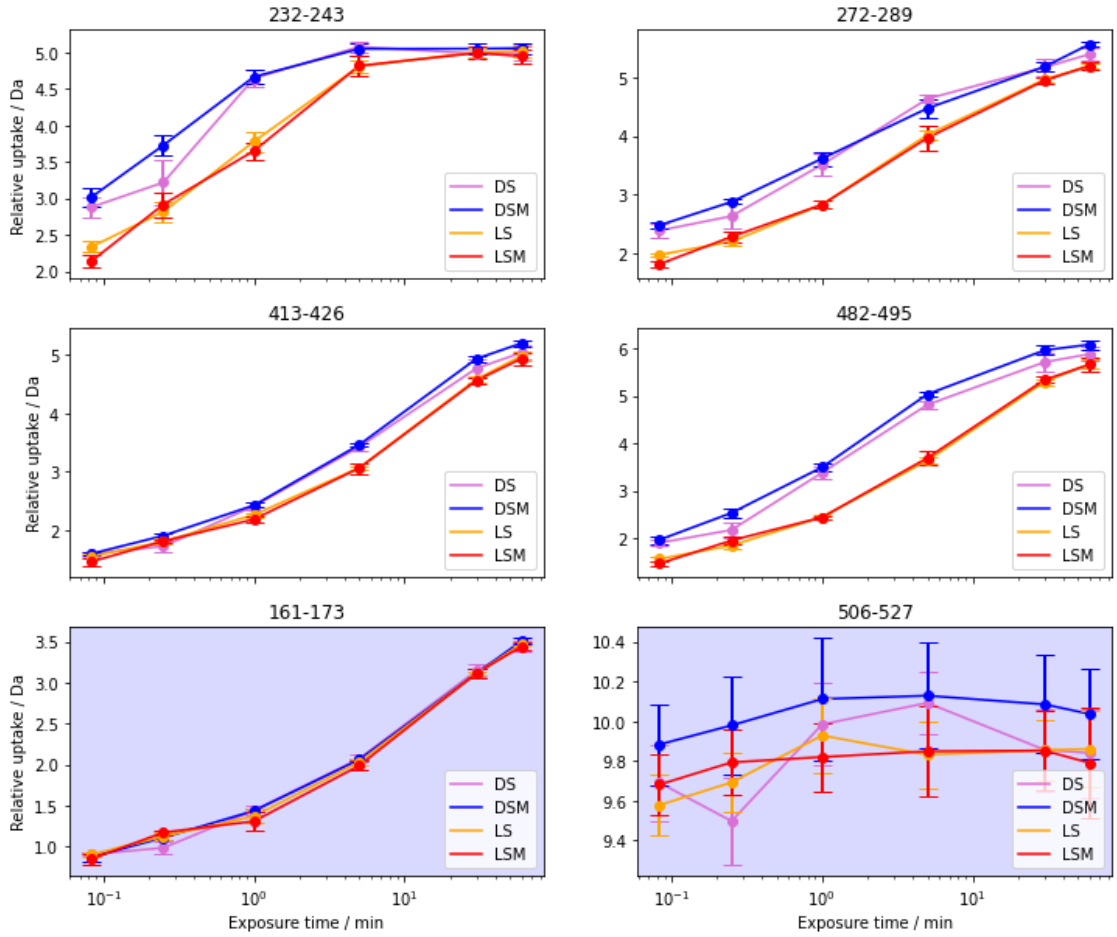

Figure 3.5: Relative uptake of representative *Er*Cry4a W369F peptides. Peptide uptake curves and standard deviations for deuterium incorporation of six representative peptides based on results displayed in Table 3.1. The first four relative uptake plots are representative of regions in which large changes between states were observed. The relative uptake plots highlighted in blue are examples of peptides with no or less clear uptake differences between states. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

example of a peptide covering the C-terminal region of the protein that experiences early saturation of deuteration. Peptide 161-173 shows very similar results to those in Figure 3.2 and peptide 506-527, with its high overall deuterium incorporation, points out the difficulties in interpreting the data of the highly surface accessible regions, again similar to what had been found for the WT protein. Figures 3.1A and 3.3A further stress this finding as it shows that the C-terminal region's uptake is already mostly saturated at the earliest time points. This makes the interpretation of uptake differences between different states especially difficult for this particular region.

Due to the time and efficiency reasons mentioned above, the repeats for the different states for this experiment were performed one after another without randomising the order of the runs. The experiments were performed in the order DSM, LSM, LS, and DS. Therefore, for this experiment the states including magnets had been investigated before the ones without the addition of magnets. Figure 3.5 points out that in this case all results are closer in uptake values to those of the earlier runs performed for the WT experiment with no obvious correlation of total uptake and order of experimental repeats. This lack of time correlation points towards the W369F mutant being a more stable protein than its WT counterpart.

Table 3.1: *Er*Cry4a W369F peptides displaying relative uptake differences between light and dark states. Peptides were only chosen if differences were present for at least three out of six time points with at most one repeat leading to an overlap of standard deviations between a dark and a light state. Only peptides that already showed a difference in uptake between LS and DS were analysed for LSM-LS and DSM-DS differences. These differences were only listed if they were present for at least two subsequent time points with a difference of at least 0.05 Da between the average values of two states. Overlap of standard deviations was not considered in this case. Overlapping peptides were defined as overlapping in two or more residues. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

Peptides showing light-induced changes	**Overlapping peptides that do not display light-induced differences**	**Differences in relative uptake upon addition of a magnet**	**Time points affected by magnet**
24-36	8/9 (22-31, 23-29, 23-31, 25-29, 26-31, 30-35, 32-37, 33-37)	DSM > DS	30 min, 60 min
100-109	8/9 (100-106, 100-112, 100-107, 101-113, 101-112, 101-107, 102-113, 108-113)	-	-

Continued on next page

Table 3.1 – continued from previous page

Peptides	**Overlapping peptides that do not display differences**	**Differences in relative uptake visible upon addition of a magnet**	**Time points affected by magnet**
129-142	7/8 (124-133, 124-136, 127-137, 127-136, 127-140, 134-140, 137-150)	DSM > DS	5 min, 30 min, 60 min
161-172	5/6 (161-173, 161-174, 161-175, 167-175, 171-192)	DS > DSM, LS > LSM	5 s, 15 s, 1 min, 5 min, 30 min, 60 min
199-216	1/2 (199-219)	DSM > DS	5 min, 30 min, 60 min
216-227	2/4 (220-227, 222-227)	DSM > DS	30 min, 60 min
217-227	2/4 (220-227, 222-227)	DSM > DS	30 min, 60 min
232-243	-	-	-
265-270	-	DSM > DS	30 min, 60 min
265-272	-	-	-
265-284	-	DSM > DS	5 s, 15 s, 30 min, 60 min
271-284	-	DSM > DS	5 s, 15 s, 30 min, 60 min
271-289	-	DSM > DS	5 s, 15 s, 30 min, 60 min
272-289	-	DSM > DS	5 s, 15 s, 1 min
285-289	-	DSM > DS	30 min, 60 min
346-361	4/6 (343-360, 343-361, 343-362, 346-360)	DS > DSM	1 min, 5 min, 30 min
346-362	4/6 (343-360, 343-361, 343-362, 346-360)	DSM > DS	15 s, 1 min, 5 min, 30 min, 60 min
362-368	-	DSM > DS	5 min, 30 min, 60 min
362-369	-	DSM > DS	30 min, 60 min

Continued on next page

Table 3.1 – continued from previous page

Peptides	**Overlapping peptides that do not display differences**	**Differences in relative uptake visible upon addition of a magnet**	**Time points affected by magnet**
363-368	-	DSM > DS	30 min, 60 min
396-402	1/2 (396-400)	-	-
404-410	-	DSM > DS	5 s, 15 s, 1 min
404-412	-	DSM > DS	5 s, 15 s
412-426	-	DSM > DS	15 s, 1 min, 30 min, 60 min
413-425	-	DSM > DS	15 s, 30 min, 60 min
413-426	-	DSM > DS	15 s, 30 min, 60 min
417-426	-	DSM > DS	15 s, 30 min, 60 min
459-481	1/2 (451-463)	DSM > DS	5 min, 30 min, 60 min
464-481	-	DSM > DS	15 s, 5 min, 30 min, 60 min
482-495	1/4 (487-506)	DSM > DS	all time points
484-495	1/4 (487-506)	DSM > DS and LSM > LS	5 min, 30 min
496-505	1/4 (487-506)	DSM > DS	5 s, 15 s, 1 min, 30 min, 60 min

To have a comprehensive list of all peptides that experience any changes when comparing the different states with each other, Table 3.1 was compiled displaying all peptides that showed relative uptake differences between LS and DS. However, the data were not subjected to a more rigorous statistical analysis and only function to give an overview of protein regions of interest. Due to the uncertainty with respect to the data quality, further statistical testing was only applied to the experimental results discussed in chapter 4.

Only peptides with LS-DS differences that were present in at least three out of six time points were listed with at most one repeat at any time point leading to

an overlap of standard deviations. The table also mentions all visible LSM-LS and DSM-DS differences as long as at least two subsequent time points with a difference of at least 0.05 Da between the average value of the two different states could be found. For this, only peptides that were already listed due to their LS-DS differences were considered. The overlap of standard deviations was not taken into account in this case to have an exhaustive list of visible differences despite the LSM-LS and DSM-DS differences being rather small. The relative uptake curves of all peptides mentioned in the table can also be found in Figures A.2 - A.5 in the Appendix as well as in Figures A.6 - A.9, also in the Appendix, on a logarithmic time scale.

The first seven rows of Table 3.1 show LS-DS differences that are not mirrored in several overlapping peptides which reduces the significance of these findings. However, the peptides covering residues 216-227 and 217-227 overlap with each other and both show a slightly higher uptake in the DSM than in the DS. Peptide 161-172 stand out as it shows a difference between LSM and LS as well as between DSM and DS where LS and DS show a higher uptake than their respective magnet counterparts, LSM and DSM. However, as mentioned above, many overlapping peptides in the same region do not display the same differences between states. Between residues 232-289 there are several overlapping peptides that display differences between LS and DS as well as between DSM and DS for six out of those eight peptides, namely 265-270, 265-284, 271-284, 271-289, 272-289, and 285-289. For these, all DSM states show a higher uptake than their DS counterparts. Further regions within which no overlapping peptides contradict the finding of LS-DS differences are between residues 362-369 and 404-426.

Peptide 346-361 shows a higher uptake for its DS than for its DSM counterpart whereas all the other subsequently listed peptides show the opposite trend with DSM having a higher uptake than the DS. The other two exceptions are peptide 396-402, where no DSM-DS difference exists, and peptide 484-495 where both LSM and DSM have a higher uptake than their respective counterparts LS and DS. However, as mentioned above, these tentative results hint at tendencies, but are

in need of verification by further experiments, especially to probe the magnetic field-induced impact on conformational changes.

Similar results for further experiments discussed in chapter 4 have additionally been subjected to a much more rigorous statistical analysis. This has not been done for the results in this chapter due to the uncertainty with respect to the data quality.

Finally, a sample quality check using native MS shows that both the *Er*Cry4a WT and the *Er*Cry4a W369F samples are mostly monomeric with some dimer being present when analysed at a concentration of approximately 8 µM. From Figure 3.6 it is evident that the monomer to dimer ratio is lower for the two WT samples (both used for the same experiment as the only difference was a slightly different purification method) than for the W369F sample which could be another reason for why the *Er*Cry4a W369F data seems to be more reliable.

3.3.1 Could different buffer conditions improve our results?

As a first means to try and improve the stability of the proteins, a comparison of different buffer conditions was performed. For this purpose, quail or *Coturnix coturnix* cryptochrome 4a (*Cc*Cry4a WT) was chosen as a cryptochrome species from another migratory bird. The reasons for this being availability from the protein producing collaborating lab and ease of handling of this particular protein. Since the two protein species have a 93.6 % sequence similarity (as calculated using the EMBOSS Needle Pairwise Sequence Alignment tool [21]), a similar behaviour under the same buffer conditions would be expected. Another cryptochrome species, chicken or *Gallus gallus* cryptochrome 4a (*Gg*Cry4a) WT, taken from a non-migratory bird, was available at the same time and planned to also be analysed at different buffer conditions to increase the certainty that the data acquired yields information on general protein stability. However, the data quality for this protein did not meet high enough standards to be used for further analysis and the results are therefore not included in this thesis. However, both *Cc*Cry4a WT and *Gg*Cry4a WT are investigated and discussed in further detail in chapter 4.

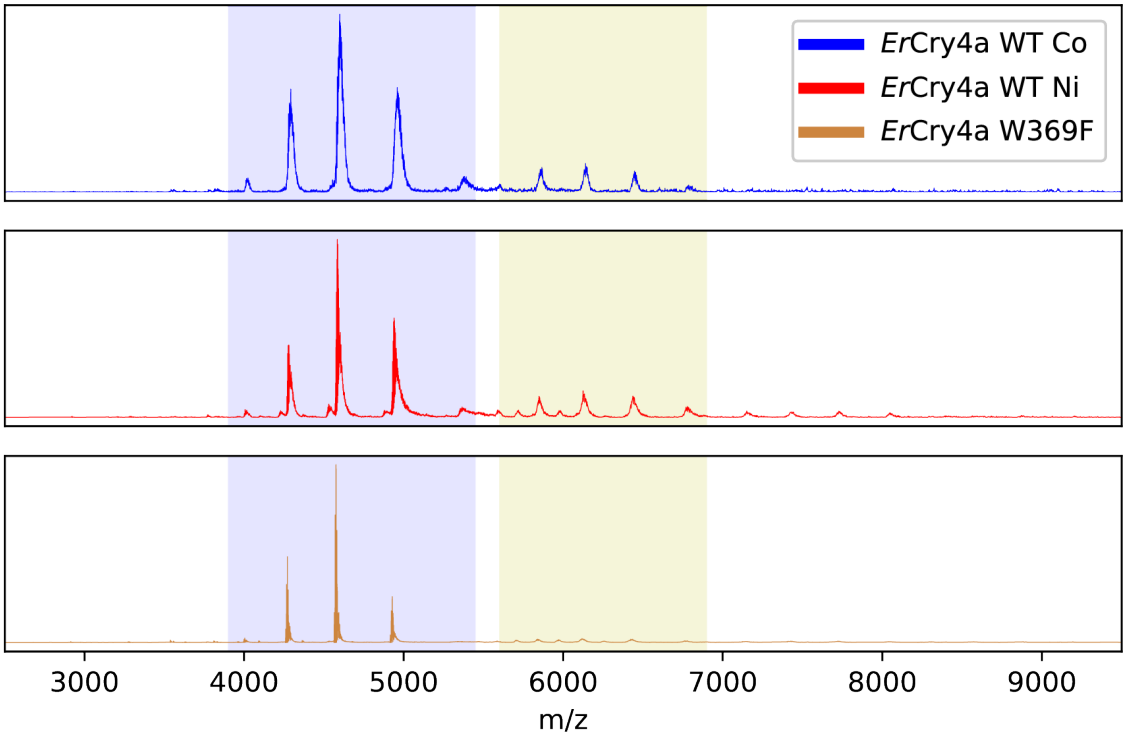

Figure 3.6: Native mass spectra of *Er*Cry4a WT and *Er*Cry4a W369F. *Er*Cry4a WT and *Er*Cry4a W369F mass spectra with a monomeric charge state distribution visible between ~3900-5500 m/z (highlighted in blue) and a dimeric charge state distribution visible between ~5600-6900 m/z (highlighted in yellow). One *Er*Cry4a WT mass spectrum had been taken of a sample after elution from a Cobalt column, the other one after elution from a Nickel column as indicated in the legend. The difference in purification procedure in the producing laboratory in Oldenburg arose due to ongoing optimization procedures. The measurements were performed on an Orbitrap Q Exactive Plus mass spectrometer (Thermo Scientific) modified for transmission of high mass species using nanoelectrospray ionisation.

The two different buffer conditions sampled are a comparison of proteins that were sent and kept in Tris or HEPES buffer throughout the duration of a standard bottom-up HDX-MS experiment as performed for the data presented above for *Er*Cry4a, with all other buffer components being the same. At the time, almost all protein samples were sent in Tris buffer. However, prior to handling of bird cryptochromes, experiments were performed on *Dm*Cry in HEPES buffer in the same laboratory which exhibited a heightened stability in comparison to the bird cryptochromes. Therefore, these two buffer conditions were chosen as a first step to compare sample stability on the time scales of hours typical for HDX-MS experiments.

The results in the form of a few representative peptides are presented in Figure 3.7 covering similar regions as in Figures 3.2 and 3.5 above. However, a comparison of differences between states can not be made as the results presented in Figure 3.7 only cover two out of the six labelling times displayed in the other two figures.

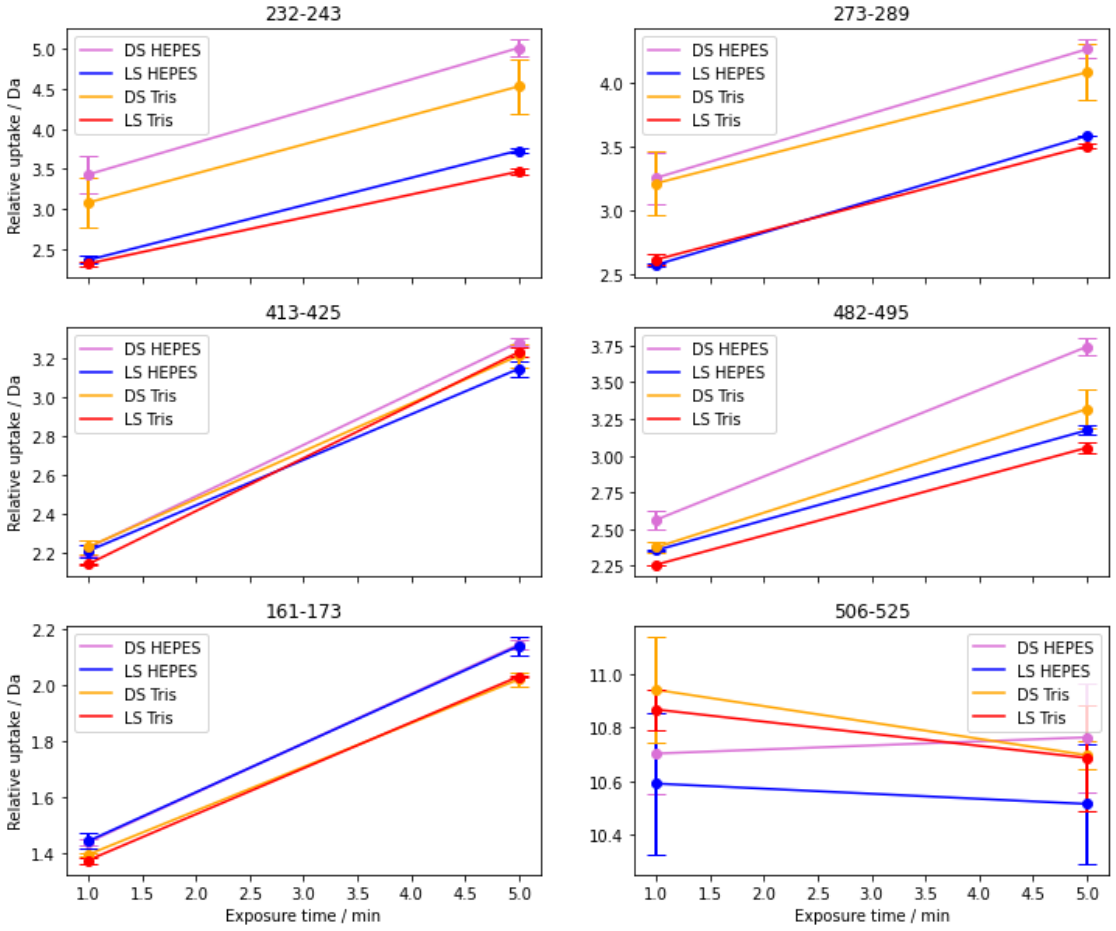

Figure 3.7: Relative uptake of representative *Cc*Cry4a WT peptides at different buffer conditions. Peptide uptake curves and standard deviations for deuterium incorporation of representative peptides for comparison of the influence of different buffer conditions on data quality. The peptides were chosen to cover similar regions as those displayed in Figures 3.2 and 3.5 while covering fewer time points. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

The differences between standard deviations as a measure for stability for all displayed peptides between the same states with different buffer conditions are rather marginal with only the DS runs in Tris buffer sometimes displaying larger error bars than other states. Where differences between light and dark states are visible, the

same states follow similar trends irrespective of the buffer conditions. One notable difference is that the overall uptake of the states performed while using HEPES buffer is higher than when the proteins were kept in Tris buffer. As this difference does not seem to impact on the differences between different states, namely general uptake trends and standard deviation, the conclusion is that the use of HEPES buffer does not significantly change the protein stability, for the better or for the worse.

Therefore, we instead resorted to improving the protein handling at the HDX-MS setup by altering the workflow to increase the speed with which the experiments were performed and decrease the time spent on ice and at room temperature for the proteins. Aside from the two buffer conditions presented here, additional different buffer conditions were also sampled on other experimental setups by collaborators and colleagues, the results of which are not presented in this work. These changes to the buffer conditions, including different concentrations of sodium chloride, the addition of trehalose, and the replacement of the reducing agent *beta*-mercaptoethanol with dithiothreitol, did not seem to improve the protein stability sufficiently on short time scales to justify sampling these on the HDX-MS setup, being a method of rather slow experimental pace and prone to frequent instrumental ailments.

3.3.2 Tracking down the C-terminal tail

As mentioned above, it seemed rather difficult to analyse the changes occurring in the C-terminal region of *Er*Cry4a with similar results for *Cc*Cry4a and *Gg*Cry4a (see chapter 4 for more details). As a means to approach this problem, we used different concentrations of deuterated water in the labelling buffer. Lower concentrations of deuterated water lead to a reduced expansion of the isotopic peptide envelope which in turn increases the signal intensity and improves ease of differentiation from other peptides [22]. This was hoped to increase the quality of the data acquired for the elusive C-terminal region and to enable differentiation between potentially different degrees of uptake in the different states probed. Again, due to availability and ease of handling, we chose *Cc*Cry4a WT as our trial protein. For this experimental

approach, we also used a newly available sample handling robot. The work presented in this section was a collaborative effort together with Laila M. N. Shah, another PhD student working in the research group of Justin L. P. Benesch.

The labelling introduced via HDX is transient [23] and the time some commercially available liquid handling robots take to prepare the labelled proteins for injection into the HDX manager for digestion and analysis might be too long to prevent back-exchange efficiently despite carefully controlled quenching conditions. This is a problem well known within the HDX-MS community and is often observed at labelling time points of one minute or longer as the relative uptake tends to go down for a lot of peptides. However, the newer version of the commercially available LEAP robot (LEAP HDX-2 Automation Manager, Waters), which was acquired during the work on this project, reduces this problem to such an extent that the robot runs start to approximate manually performed runs. The relative uptake for longer time points does not tend to decrease anymore. However, as visible in Figure 3.8A for peptide 506-525 as a representative peptide for the whole C-terminal region, this problem still seems to occur for those peptides which are highly solvent exposed and saturated at early labelling time points. The C-terminal region starting at approximately residue 500 seems to be sufficiently solvent accessible so that it is already saturated during early time points for the labelling solutions of lower D_2O concentrations, similar to when the solution is exposed to 92 % of deuterated water. Therefore, the usage of different concentrations of deuterated water does not seem to lead to a lot more information on the C-terminal region other than that the region is very easily saturated and therefore very solvent exposed and/or very flexible.

Peptide 232-243 shows that a concentration of 50 % deuterated water already leads to visible differences in relative uptake between the LS and the DS with the DS being saturated after 1 min of labelling time while the LS's uptake is still increasing after this time point. The uptake differences between LS and DS at 10 % D_2O are very small in comparison to the other percentages of deuterated water that were used. Initially, 75 % of D_2O was also applied, but the results were very similar to those obtained with a concentration of 92 % and therefore not further investigated.

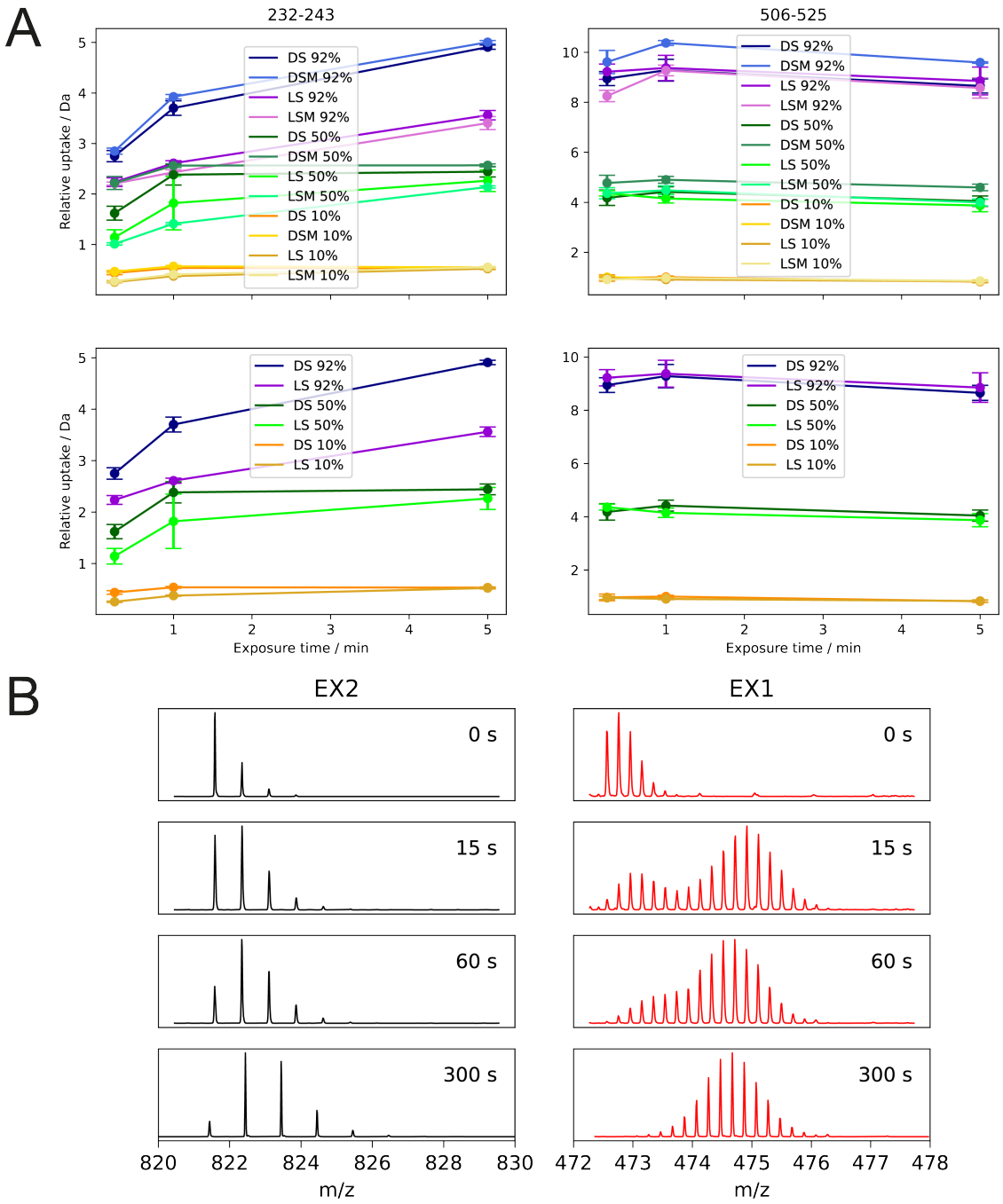

Figure 3.8: Results of a D_2O concentration trial. (**A**) Peptide uptake curves and standard deviations for deuterium incorporation of representative *Cc*Cry4a WT peptides. First row: Peptides 232-243 and 506-525 after incubation with three different concentrations of deuterated water including the states DS, DSM, LS, and LSM. Second row: Same figures as in first row excluding information on the states with magnets for easier visualisation. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet. (**B**) Example of EX1 behaviour found in the C-terminal tail region (exemplified using peptide 506-525, charge state +5) in comparison to the more common EX2 behaviour (exemplified using peptide 362-368, charge state +1). Further analysis of this data set has been performed by Laila M. N. Shah and will not be covered within this thesis but instead discussed elsewhere.

Figure 3.8B points towards another finding after acquiring this large data set: the isotope distributions of some of the peptides covering the C-terminal region exhibit EX1 behaviour. The in depth analysis of the full data set was conducted by Laila M. N. Shah and will be presented elsewhere. This finding is best visible in the data acquired with the highest percentage of deuterated water used. Slysz *et al.* mention in their work on the effects of different deuterium concentrations for HDX-MS that higher concentrations leading to a more pronounced isotopic envelope might be required to thoroughly investigate this behaviour [22]. The presence of EX1 behaviour might indicate slow opening or closing event within this region where parts of the proteins in solution are more solvent accessible than others within the C-terminal region.

3.3.3 Concluding discussion

This chapter acts as an introduction to the investigation of avian cryptochromes using HDX-MS in order to ease the reader into the results presented in chapter 4. The successful data acquisition and analysis for a large comparative study of different avian cryptochromes presented therein builds on the results presented in this chapter.

Light-induced conformational changes are clearly visible in several peptides across the sequence of *Er*Cry4a W369F which mostly display protection events. Due to the difficulties encountered in obtaining reliable data from *Er*Cry4a WT and the resulting uncertainty about how trustworthy the results obtained for *Er*Cry4a W369F are, also given they were obtained under exactly the same problematic experimental conditions (keeping the samples at room temperature for several hours without randomising the states during the experiment), all data presented in this chapter should be considered with caution. However, the results can be taken as a guide and illustration of first insights into conformational changes of avian cryptochromes.

Although differences between LS and DS have been consistently found in proteins from different birds, the LSM-LS and DSM-DS differences are less obvious and in need of statistical validation (with the exception of *Er*Cry4a WT where it is the

other way around with the reason for this being quite likely sample instability). Interestingly, the difference most repeatedly observed in the detailed analysis of the *Er*Cry4a W369F data set is an increased uptake for the DSM state in comparison to the DS. This seems unexpected given that the radical pair mechanism requires light in order to allow the perception of the inclination of the Earth's magnetic field, leading to the expectation of an additional magnetic field only influencing the LS rather than the DS. It should be noted here, that the strength of the magnets applied exceeds the one that the Earth's magnetic field exerts by about three orders of magnitude (see Figure 3.9 below for details). The Earth's magnetic field was not cancelled out during the experiments. However, this is unlikely to influence the experimental results since the additionally applied, static magnetic field is so much stronger than the Earth's magnetic field and since the literature so far does not provide experiments where a measurable magnetic field effect was the result of an Earth strength magnetic field exposure. Furthermore, as the proteins were analysed while freely tumbling around in solution while a certain degree of alignment is assumed for the radical pair mechanism to function as a detector of the direction of the Earth's magnetic field [6], this further complicates the interpretation of the results. If the additional magnetic field has an influence on the conformational changes, it might have a slightly varied effect on the proteins depending on their initial orientation in solution. However, since all measurements were performed after at least 20 min of incubation in their respective states (LS, DS, etc.), it would be conceivable to argue that the time frame could be sufficient for all proteins to be aligned in the same way with an externally applied magnetic field at some point during the incubation period. It is likely that the protein's radical lifetimes are longer than the time they need to tumble around in solution given a previously calculated rotational correlation time of *Er*Cry4a WT of about 80 ns (in 20% glycerol and at 5°C) [24]. Additionally, the isotropic magnetic field effects are likely to be larger than the anisotropic ones [25], making it less likely for anisotropy to play a role in the results obtained.

The findings regarding the C-terminal tail, given that the current literature gives a lot of attention to this part of the protein assuming it to be an important part for signal transduction [14, 26–29], seems to be a bit discouraging at first sight. The use of different deuterium concentrations during the labelling process does not seem to increase the information on a mechanism by which the C-terminal tail might act following the internal radical formation. However, different approaches might yield more information such as adding shorter labelling times (see chapter 4). In addition to this, Shah's in depth analysis of the EX1 behaviour in this region might yield further interesting results.

To probe how to increase protein stability over time, the main buffer component was changed from Tris to HEPES. However, this does not seem to have a noteworthy impact. Since the proteins were kept on ice for several days and at room temperature for several hours during a single day, we resolved to greatly reduce the time the proteins spend in either of these conditions by asking our collaborators who produce the proteins to send them in several aliquots, reducing the amount of protein that was thawed and kept on ice at any one time without having to introduce several thawing and re-freezing steps as these seem to further decrease protein stability. We also decided to reduce the time spent at room temperature down to a maximum of three hours for future experiments. Both these changes allow for more flexibility in the experimental design, leaving more time in between different runs where the setup can be changed (e.g. to include the magnets) without prolonging the time the proteins spend on the setup even further. Additionally, randomising the order of the states and labelling times is generally recommended for HDX-MS experiments [30] and would therefore lead to more reliable data sets.

3.3.4 Final remarks

In this chapter, we have found that HDX-MS is a suitable technique to investigate avian cryptochromes. However, a number of changes to the experimental procedures need to be implemented in order to increase protein stability, reproducibility,

and confidence in our findings. These have been implemented for the work presented in chapter 4.

Two additional difficulties have been encountered during the work presented in this chapter, which is the sample availability in connection with HDX-MS setup availability. Due to a large number of people working with the same proteins from the same collaborating lab, compromises had to be made. This is why some of the experiments aimed at improving protein stability could not be tested with the seemingly least stable protein *Er*Cry4a WT and the results obtained using other proteins might not be applicable to the same extent to this particular protein. Furthermore, the HDX-MS setup available for this work is shared between multiple users and experienced frequent break-downs due to a diverse range of reasons requiring more time spent on fixing the instrument rather than performing experiments. Since a single experiment presented herein can take up to about 80 hours of experimental time and since experience tells that avian cryptochromes might not uphold their full structural integrity for longer than 1.5-2 months, frequent instrument breakdowns greatly slow down the overall progress and limit the exploration of possible changes to the setup and workflow.

3.4 Material and Methods

3.4.1 Material

Reagents and chemicals were purchased from Sigma Aldrich (Gillingham, UK). Proteins were expressed in and purified from *E. coli* by our collaborators in the group of Henrik Mouritsen at the Carl von Ossietzky University of Oldenburg, Germany. The purification and expression was performed as described before [7]. Unless indicated otherwise, all samples were shipped on ice in 20 mM Tris, 250 mM NaCl, 2 mM DTT, and 20% glycerol and kept at -80°C until further use.

The equilibration buffer consisted of 5 mM potassium phosphate (H_2O), pH 7.0, the labelling buffer consisted of 5 mM potassium phosphate (D_2O), pD 7.0, and the quench buffer consisted of 50 mM potassium phosphate (H_2O), pH 1.9.

3.4.2 HDX-MS

Essential protein handling at the HDX-MS setup was performed under dim red light conditions in an otherwise completely dark room. The proteins were diluted down to 10-16 µL prior to the start of the experiment.

All proteins were equilibrated for at least 20 min in a temperature controlled chamber at 20°C under different environmental conditions prior to labelling: the light state of the protein was induced by continuous illumination with blue LEDs at 450 nm and powers between 0.03-0.06 mW cm^{-1}, the dark state of the protein was kept in complete darkness and the magnet states were induced by inserting neodymium magnets (10 mm x 1.5 mm N42 neodymium magnets with 1 kg pull; Magnet Expert Ltd.) of 160 mT strength below the samples in a custom-built vial holder. The magnet strength directly applied to the proteins lies between ∼18-25 mT with a distance of approximately ∼2-3 mm from the magnet as shown in Figure 3.9.

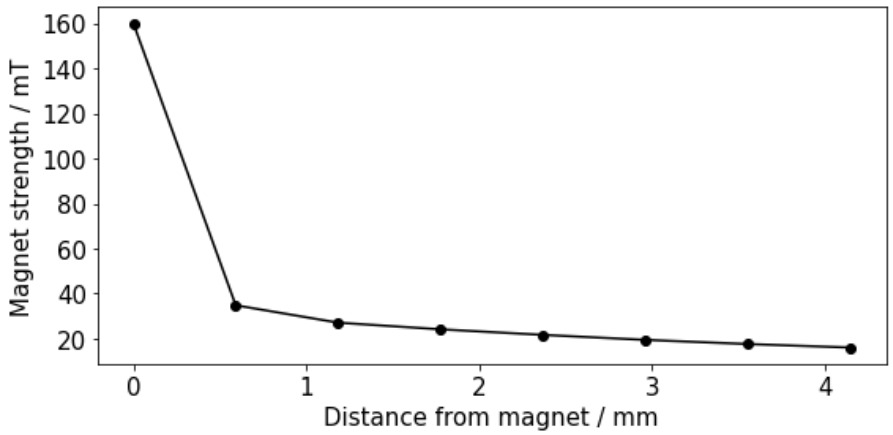

Figure 3.9: Magnet strength dependence on distance. Magnet strengths as measured with a gaussmeter by Patrick D. F. Murton.

Proteins were labelled by adding 55 µL of the labelling buffer at room temperature to 5 µL of sample resulting in a labelling solution of ∼92% D_2O for *Er*Cry4a W369F. For *Er*Cry4a WT, 54 µL of the labelling buffer at room temperature were added to 1 µL of sample resulting in a labelling solution of ∼98% D_2O. The latter approach was abandoned for the W369F sample as aliquoting only 1 µL seemed to bear the risk of inaccuracy due to residual liquid on the outside of the pipette tip. After labelling,

the reactions were quenched by adding an equal volume of a quench solution that had been kept at 0°C followed by immediate injection of the quenched solution into the HDX manager (Waters, Milford, MA, USA). Subsequently, the solution was kept at 0°C until reaching the pepsin column Enzymate Pepsin (Waters) for digestion where it was held at 20°C using an isocratic H_2O (0.1 % v/v) formic acid solution at a flow rate of 100 µL/min. This was followed by 2 min of product collection and de-salting on a BEH C_{18} trapping column (Waters) and subsequent separation on a reversed-phase analytical column (Waters) at 0°C using a flow rate of 40 µL/min.

The elution gradient for separation started with 97 % H_2O and 3 % MeCN changing to 65 % H_2O and 35 % MeCN within 7 minutes followed by a gradient going from 65% to 5% H_2O within minutes 7-8 and held at 5% H_2O for minutes 8-10. All peptides were electrosprayed into a Synapt G2Si Q-ToF instrument (Waters) for mass analysis and all sample handling was performed manually. The instrument parameters were as follows: capillary at 2.8 kV, sample cone at 100 V, source offset at 30 V, source temperature at 80°C, and the desolvation temperature at 150°C. Leucine enkephalin (LeuEnk) was used for internal calibration and its m/z acquired every 30 s.

A cleaning cycle with a saw-tooth gradient was performed in between each labelling protein run using a washing solution of 4 % acetonitrile (ACN) and 0.1 % formic acid in water with 0.5 M guanidinium hydrochloride (GdnHCl). Intermittent blank runs were injected to ensure the carryover between different protein runs to be minimal.

The small red lights used to illuminate the work bench and the blue LEDs used to illuminate the samples when necessary were tested for their emission spectra to avoid unwanted overlap of wavelengths. The recorded emission spectra displayed in Figure 3.10 have been acquired by Patrick D. F. Murton using a spectrograph and a CCD camera. The same lights were used for all experiments presented in this thesis.

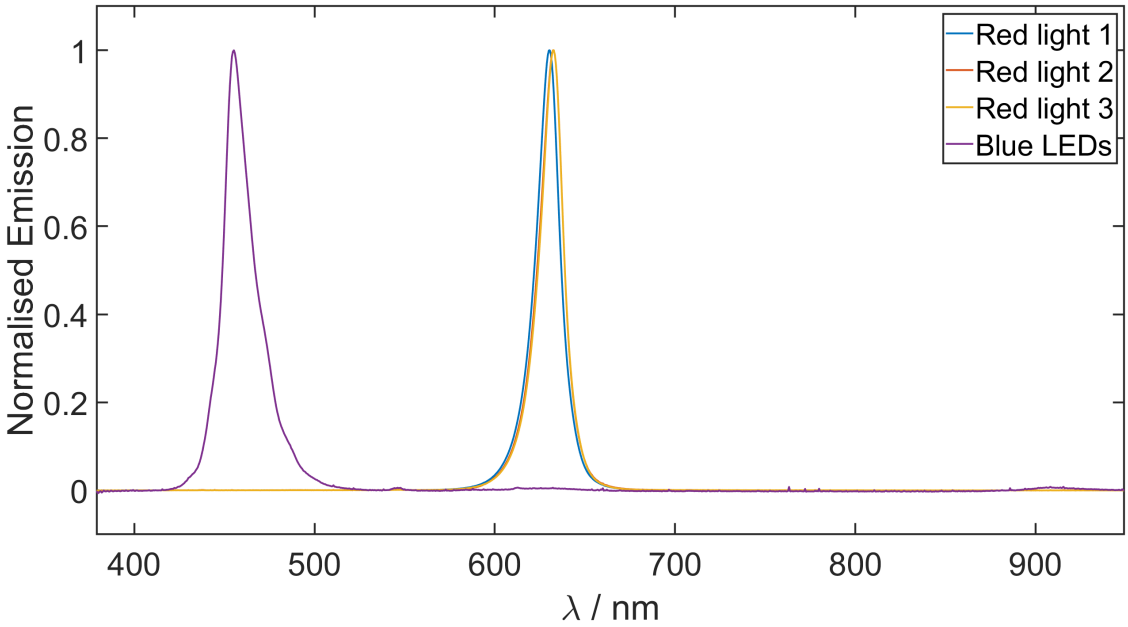

Figure 3.10: Emission spectra of all lights used during experiments.

3.4.3 Native MS

The *Er*Cry4a samples used for these experiments were the same as used for the HDX-MS experiments. The samples were buffer-exchanged into 200 mM ammonium acetate, pH 8, for the native MS measurements using Zeba^{TM} Micro Spin desalting columns with a molecular weight exclusion limit of 40 kDa (ThermoFisher Scientific).

Measurements were performed in Oxford using in-house gold-plated capillaries on a Q Exactiv^{TM} mass spectrometer in positive ion mode. The samples were sprayed at approximately 8 µM using nanoflow electrospray with a source temperature of 150°C and a capillary voltage of 1.0 kV [31]. The higher-energy C-trap dissociation (HCD) cell voltage was 5 V, in-source trapping was set to -200 V to help with the dissociation of small ion adducts, and the noise threshold parameter was 3. Ion transfer optics and voltage gradients throughout the instruments were adjusted for optimum protein transmission.

3.4.4 Data analysis

Peptides were identified using MS/MS analysis (MS^E) on equilibrated, but unlabelled proteins and processed using the Protein Lynx Global server 3.0 software (Waters)

for database searching. They were subsequently read into the DynamX 3.0 software (Waters), an automated peak assignment tool. Only peptides passing the applied fragmentation quality filters of a minimum fragmentation products per amino acid of 0.2, a mass accuracy with a maximum $[MH]^+$ error of 5 ppm, a maximum sequence length of 25 amino acids, and a minimum reproducibility defined by peptides identified in at least three MS^E repeats, were used for further analysis. All labelled peptides were mapped onto this list of peptides. The automated peak assignment was verified for each charge state for each peptide with overlapping or noisy spectra being discarded. The deuterium content of the labelled samples was determined by comparison to the non-deuterated control sample to determine the relative m/z shift.

Native mass spectra were processed using the MassLynx V4.2 software (Waters, Wilmslow, UK).

The PyMOL software [32] (Schrödinger, NY, USA) was used for 3D data visualisation.

References

1. Parak, F. Proteins in action: the physics of structural fluctuations and conformational changes. *Current Opinion in Structural Biology* **13,** 552–557. https://linkinghub.elsevier.com/retrieve/pii/S0959440X03001374 (2003).
2. Wiltschko, W. & Merkel, F. W. Orientierung zugunruhiger Rotkehlchen im statischen Magnetfeld. *Verh. Dtsch. Zool. Ges* **59,** 362–367.
3. Wiltschko, W. & Wiltschko, R. Magnetic Compass of European Robins. *Science* **176,** 62–64. https://www.science.org/doi/10.1126/science.176.4030.62 (1972).
4. Wiltschko, W. & Wiltschko, R. The effect of yellow and blue light on magnetic compass orientation in European robins, Erithacus rubecula. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology* **184,** 295–299. http://link.springer.com/10.1007/s003590050327 (1999).
5. Mouritsen, H. in *Sturkie's Avian Physiology* 113–133 (Elsevier, 2015). https://linkinghub.elsevier.com/retrieve/pii/B9780124071605000087.
6. Hore, P. J. & Mouritsen, H. The Radical-Pair Mechanism of Magnetoreception. *Annual Review of Biophysics* **45,** 299–344 (2016).

7. Xu, J. *et al.* Magnetic sensitivity of cryptochrome 4 from a migratory songbird. *Nature* **594,** 535–540. http://dx.doi.org/10.1038/s41586-021-03618-9 (2021).
8. Görtemaker, K. *et al.* Direct Interaction of Avian Cryptochrome 4 with a Cone Specific G-Protein. *Cells* **11,** 2043. https://www.mdpi.com/2073-4409/11/13/2043 (2022).
9. Timmer, D. *et al.* Tracking the Electron Transfer Cascade in European Robin Cryptochrome 4 Mutants. *Journal of the American Chemical Society* **145,** 11566–11578. https://arxiv.org/abs/2205.10393?utm_source=researcher_app&utm_medium=referral&utm_campaign=RESR_MRKT_Researcher_inbound%20https://pubs.acs.org/doi/10.1021/jacs.3c00442 (2023).
10. Franz-Badur, S. *et al.* Structural changes within the bifunctional cryptochrome/photolyase CraCRY upon blue light excitation. *Scientific Reports* **9,** 9896. https://www.nature.com/articles/s41598-019-45885-7 (2019).
11. Chandler, S. *Probing protein structure and dynamics through native MS and H/D exchange strategies* PhD thesis (University of Oxford, 2019).
12. Emmerich, H.-J. *et al.* A topologically distinct class of photolyases specific for UV lesions within single-stranded DNA. *Nucleic Acids Research* **48,** 12845–12857. https://academic.oup.com/nar/article/48/22/12845/6018439 (2020).
13. Eckel, M., Steinchen, W. & Batschauer, A. ATP boosts lit state formation and activity of Arabidopsis cryptochrome 2. *The Plant Journal* **96,** 389–403. https://onlinelibrary.wiley.com/doi/10.1111/tpj.14039 (2018).
14. Watari, R. *et al.* Light-dependent Structural Change of Chicken Retinal Cryptochrome4. *Journal of Biological Chemistry* **287,** 42634–42641. https://linkinghub.elsevier.com/retrieve/pii/S0021925820437561 (2012).
15. Zoltowski, B. D. *et al.* Structure of full-length Drosophila cryptochrome. *Nature* **480,** 396–399. https://www.nature.com/articles/nature10618 (2011).
16. Wales, T. E. & Engen, J. R. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrometry Reviews* **25,** 158–170. https://onlinelibrary.wiley.com/doi/10.1002/mas.20064 (2006).
17. Houde, D., Berkowitz, S. A. & Engen, J. R. The Utility of Hydrogen/Deuterium Exchange Mass Spectrometry in Biopharmaceutical Comparability Studies. *Journal of Pharmaceutical Sciences* **100,** 2071–2086. https://linkinghub.elsevier.com/retrieve/pii/S0022354915320955 (2011).
18. Lin, C., Top, D., Manahan, C. C., Young, M. W. & Crane, B. R. Circadian clock activity of cryptochrome relies on tryptophan-mediated photoreduction. *Proceedings of the National Academy of Sciences* **115,** 3822–3827. https://pnas.org/doi/full/10.1073/pnas.1719376115 (2018).

19. Wong, S. Y., Wei, Y., Mouritsen, H., Solov'yov, I. A. & Hore, P. J. Cryptochrome magnetoreception: four tryptophans could be better than three. *Journal of The Royal Society Interface* **18.** https://royalsocietypublishing.org/doi/10.1098/rsif.2021.0601 (2021).

20. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nature Methods* **19,** 679–682. https://www.nature.com/articles/s41592-022-01488-1 (2022).

21. Madeira, F. *et al.* Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research* **50,** W276–W279. https://academic.oup.com/nar/article/50/W1/W276/6567472 (2022).

22. Slysz, G. W., Percy, A. J. & Schriemer, D. C. Restraining Expansion of the Peak Envelope in H/D Exchange-MS and Its Application in Detecting Perturbations of Protein Structure/Dynamics. *Analytical Chemistry* **80,** 7004–7011. https://pubs.acs.org/doi/10.1021/ac800897q (2008).

23. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. Primary structure effects on peptide group hydrogen exchange. *Proteins: Structure, Function, and Bioinformatics* **17,** 75–86 (1993).

24. Wong, S. Y., Benjamin, P. & Hore, P. J. Magnetic field effects on radical pair reactions: estimation of B 1/2 for flavin-tryptophan radical pairs in cryptochromes. *Physical Chemistry Chemical Physics* **25,** 975–982. http://xlink.rsc.org/?DOI=D2CP03793A (2023).

25. Hiscock, H. G. *et al.* The quantum needle of the avian magnetic compass. *Proceedings of the National Academy of Sciences* **113,** 4634–4639. https://pnas.org/doi/full/10.1073/pnas.1600341113 (Apr. 2016).

26. Partch, C. L., Clarkson, M. W., Özgür, S., Lee, A. L. & Sancar, A. Role of Structural Plasticity in Signal Transduction by the Cryptochrome Blue-Light Photoreceptor. *Biochemistry* **44,** 3795–3805. https://pubs.acs.org/doi/10.1021/bi047545g (2005).

27. Park, H. *et al.* Optogenetic protein clustering through fluorescent protein tagging and extension of CRY2. *Nature Communications* **8,** 30. http://dx.doi.org/10.1038/s41467-017-00060-2%20http://www.nature.com/articles/s41467-017-00060-2%20https://www.nature.com/articles/s41467-017-00060-2 (2017).

28. Parico, G. C. G. & Partch, C. L. The tail of cryptochromes: an intrinsically disordered cog within the mammalian circadian clock. *Cell Communication and Signaling* **18,** 182. https://biosignaling.biomedcentral.com/articles/10.1186/s12964-020-00665-z (2020).

29. Bradlaugh, A. A. *et al.* Essential elements of radical pair magnetosensitivity in Drosophila. *Nature* **615,** 111–116. https://www.biorxiv.org/content/10.1101/2021.10.29.466426v1%0Ahttps://www.biorxiv.org/content/10.1101/2021.10.29.466426v1.abstract%20https://www.nature.com/articles/s41586-023-05735-z (2023).

30. Masson, G. R. *et al.* Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nature Methods* **16,** 595–602. https://www.nature.com/articles/s41592-019-0459-y (2019).

31. Kondrat, D. L. F., Struwe, W. B. & Benesch, J. L. P. *Structural Proteomics* Second Edi (ed Owens, R. J.) 349–371. https://link.springer.com/10.1007/978-1-4939-2230-7 (Springer New York, New York, NY, 2015).

32. Schrödinger, L. & DeLano, W. *The PyMOL Molecular Graphics System* 2022. http://www.pymol.org/pymol.

Under the surface
The ship doesn't swerve when it heard how big the iceberg is
Under the surface
I think about my purpose, can I somehow preserve this?

— from Amanda Palmer's *Surface Pressure*

4

A comparative study of conformational changes within different avian cryptochrome species

Contents

4.1 Introduction 120
4.2 Light-induced conformational changes of avian cryptochromes 122
4.2.1 Three different avian species 122
4.2.2 Uptake comparison 125
4.2.3 Statistical data analysis 130
4.2.4 Conclusions 164
4.3 Outlook and Interaction partner studies 167
4.4 Materials and Methods 169
4.4.1 Materials 169
4.4.2 HDX-MS 169
4.4.3 Native MS 170
4.4.4 Data analysis 171
References 171

4.1 Introduction

Chapter 3 presented unexpected results using HDX-MS to analyse the cryptochrome protein *Er*Cry4a WT and its W369F mutant with a shortened electron transfer pathway. The results pointed out experimental problems which were subsequently fixed and led to reliable data sets of different avian cryptochromes and their W369F mutants. The results of these data sets are being presented in this chapter and currently also implemented in a manuscript draft.

As this chapter used very similar but improved methods and setups as described in chapter 3, some results and findings of the previous chapter apply in a similar way to this chapter and will be referenced accordingly.

The chapter contains successful and reliable repeats of the *Er*Cry4a WT and *Er*Cry4a W3609F HDX-MS experiments attempted in chapter 3 as well as the same experiments performed using quail cryptochrome 4 (*Cc*Cry4a WT), chicken cryptochrome 4 (*Gg*Cry4a WT) and both their W369F mutants. Additionally, to make sure that the light-induced changes in uptake are due to the presence of the cofactor FAD and FAD-dependent radical formation, we have performed the same experiment on the *Er*Cry4a R356E mutant as a null comparison. For this mutant, our collaborators in Henrik Mouritsen's research group at the Carl von Ossietzky University of Oldenburg introduced a change of charge near the FAD cofactor resulting in a folded protein which does not incorporate the cofactor.

As one HDX-MS experiment already provides a wealth of information and here we are presenting the results of seven of these investigating four different states, several time points and four repeats for each time point, this chapter will start by briefly introducing the protein species that were investigated followed by a step by step analysis of the different regions that displayed significant uptake differences between the different states. Each region will then be analysed separately by comparing the results of all seven experiments for this particular region, aided by multiple figures (partially in the Appendix due to their size) and tables combining the findings. Different statistical analysis approaches have been deployed and their results are presented one after another.

The outlook section will introduce potential future explorations of protein function as well as mention a few experiments that have already been performed by Laila M. N. Shah and which will complement the results presented in this chapter and feature in a publication that is currently being drafted.

A part of the data displayed in this chapter was recorded by Shah (parts of the HDX-MS experiments on *Er*Cry4a R356E, *Cc*Cry4a W369F, and *Gg*Cry4a W369F). The data analysis was performed by both the author and Shah and figures created by Shah are indicated. As such, the interpretation and analysis within this chapter have been a joint effort and are currently being combined and prepared for publication.

Originally, the aim was to complement the data sets presented in this chapter by interaction partner studies. Since six putative interaction partners of *Er*Cry4a were found by Wu *et al.* in 2020 [1], the interest in how they might interact grew. Görtemaker *et al.* were able to express one of these six candidates, the α-subunit of a cone-specific heterotrimeric G protein called guanine nucleotide-binding protein G(t) subunit α-2 (GNAT2) and found a strong interaction between this protein and *Er*Cry4a [2]. The plan was to firstly involve taking native MS spectra to verify an interaction in solution and to secondly submit the protein complex to the HDX-MS setup in order to elucidate the interaction surface between GNAT2 and *Er*Cry4a WT. Unfortunately, due to a range of different reasons (packages getting stuck in customs at the border, samples getting lost during shipment, a changeover in PhD students at the collaborating laboratory, and instrumental failures of the HDX-MS setup), progress was slowed down significantly. A few initial results will briefly be discussed in the outlook section of this chapter and the subsequent HDX-MS experiments will be carried out by Shah, who is taking over this project, in the coming weeks and months.

4.2 Light-induced conformational changes of avian cryptochromes

4.2.1 Three different avian species

As introduced above, the protein species *Er*Cry4a WT, *Er*Cry4a W369F, *Er*Cry4a R356E, *Cc*Cry4a WT, *Cc*Cry4a W369F, *Gg*Cry4a WT, and *Gg*Cry4a W369F have been investigated using HDX-MS at concentrations in the range 10-16 µM. For clarity, all results presented in the following will investigate the proteins in the order just mentioned.

Figure 4.1A shows the same AlphaFold [3] model of *Er*Cry4a WT already displayed in chapter 3 with its regions of interest pointed out that will subsequently be discussed. The areas indicated are expected to lie in similar regions within all the different investigated species due to their sequence similarity (93.6 % for robin and quail, 93.6 % for robin and chicken, and 97.2 % for quail and chicken as calculated using the EMBOSS Needle Pairwise Sequence Alignment tool [4]). Figure 4.1B displays the sequence coverages obtained for every experiment.

Alongside the HDX-MS experiments, native MS measurements of every protein that was investigated were performed at concentrations of approximately 8 µM as displayed in Figure 4.2 to assess the quality of the proteins. The apo-protein *Er*Cry4a R356E corresponds to a mass expected of this protein without containing its cofactor. All proteins were mostly monomeric with a small amount of dimer present. Some native MS spectra show a slightly larger degree of dimerisation than others and display a few higher-order oligomers such as trimers. However, it is worth mentioning here that the same proteins were subsequently diluted down for the HDX-MS experiments using a buffer containing 10 mM of BME which, as discussed in chapter 3, reduced the degree of dimerisation and higher order oligomerisation (if present). Therefore, the final experiments were performed on highly monomeric samples. Figure 4.3 corroborates this by displaying highly monomeric samples used for two of the experiments performed as analysed after three hours at room temperature in their respective state conditions (LS, DS, LSM, and DSM).

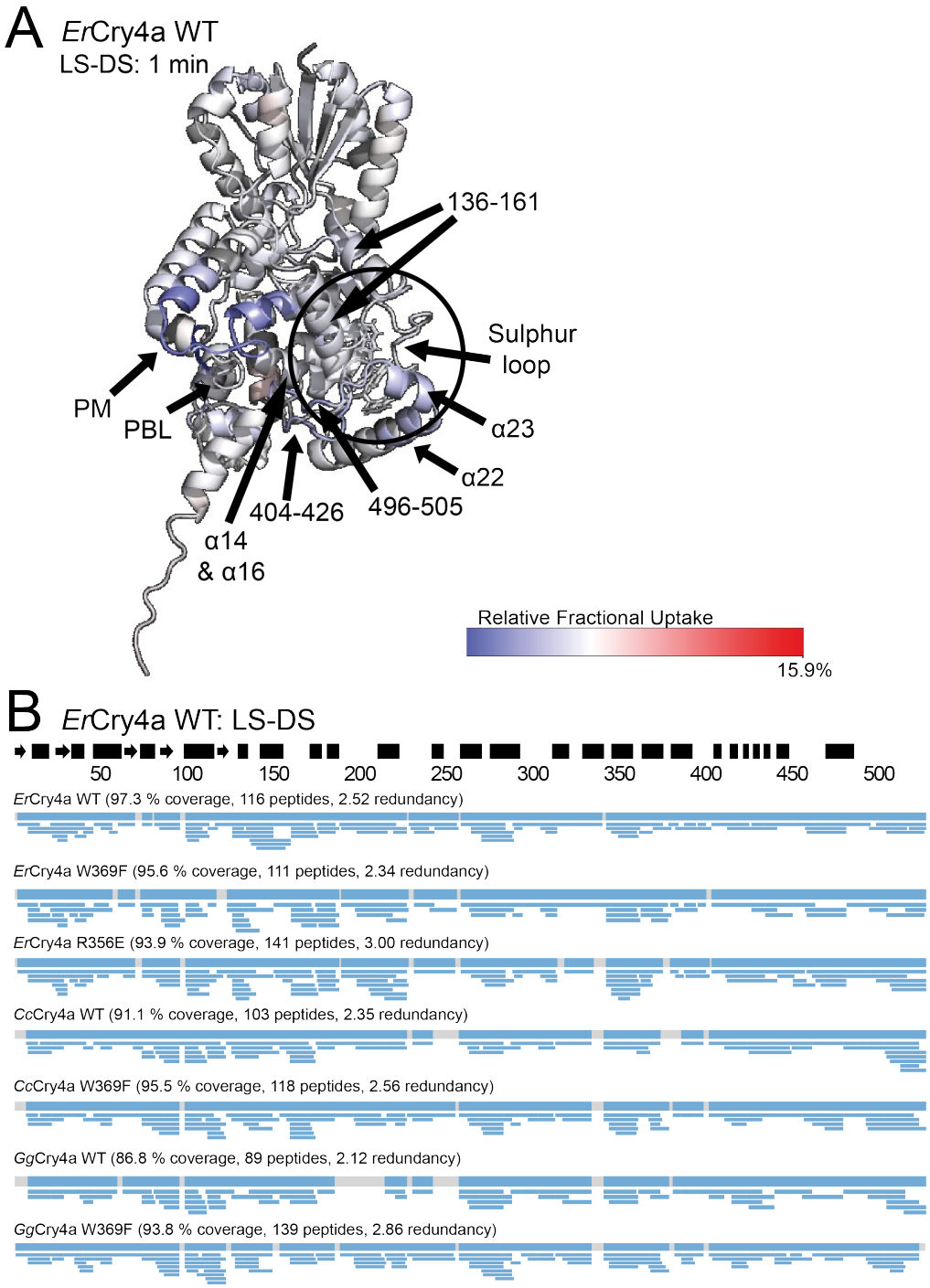

Figure 4.1: Relative fractional uptake of *Er*Cry4a WT and coverage maps. (**A**) Relative fractional uptake across the entire protein sequence after 1 min of labelling for the LS-DS comparison displayed on an AlphaFold structure. The structure was generated by Navoneel Sen using AlphaFold [3]. For regions covered by multiple overlapping peptides, the data displayed is determined by the shortest peptide. The area of the tryptophan tetrad is encircled in black. PBL - phosphate-binding loop, PM - protrusion motif. (**B**) Coverage maps of all seven proteins after statistical data analysis in Deuteros [5]. Beta sheets (arrows) and alpha-helices (rectangles) are indicated above the sequence.

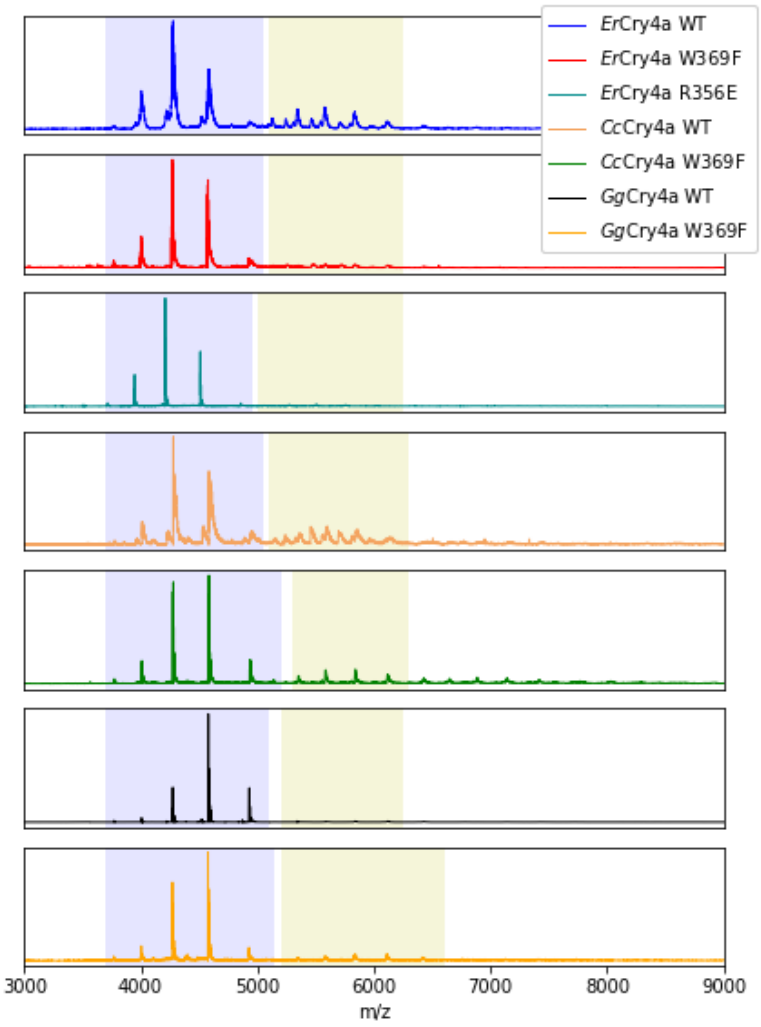

Figure 4.2: Native mass spectra of seven avian cryptochrome 4a species. For all spectra a monomeric charge state distribution is visible between ~3900-5500 m/z (highlighted in purple) and a dimeric charge state distribution is visible between ~5600-6900 m/z (highlighted in yellow). For some proteins, trimer peaks were observed at higher m/z values. The measured monomeric masses are as follows (with the theoretical masses in brackets): *Er*Cry4a WT: 64111.98±31.50 Da (64069.77 Da), *Er*Cry4a W369F: 64028.34±28.44 Da (64030.73 Da), *Er*Cry4a R356E: 63125.66±0.79 Da (63257.14 Da), *Cc*Cry4a WT: 64178.14±23.71 Da (64294.03 Da), *Cc*Cry4a W369F: 64170.34±29.84 Da (64254.99 Da), *Gg*Cry4a WT: 64088.49±0.98 Da (64219.85 Da), *Gg*Cry4a W369F: 64049.25±0.47 Da (64180.82 Da). *Er*Cry4a WT and *Cc*Cry4a WT were unfortunately only measured at a comparatively low intensity. Therefore, the actual monomer to dimer ratio might be slightly different to what is displayed above. The measurements were performed on an Orbitrap Q Exactive Plus mass spectrometer (Thermo Scientific) modified for transmission of high mass species using nanoelectrospray ionisation.

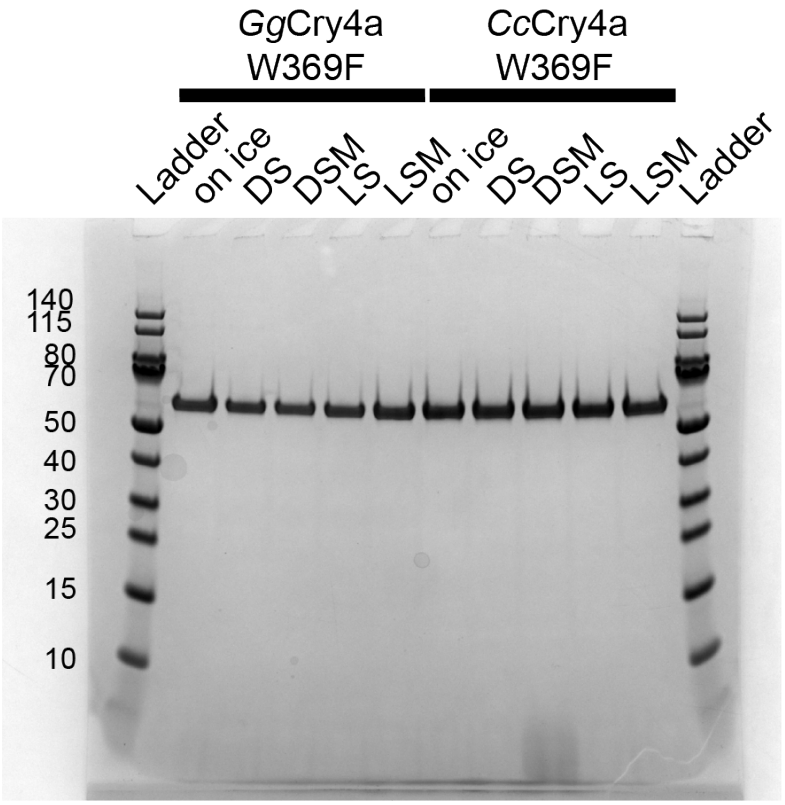

Figure 4.3: SDS-PAGE gel of *Gg*Cry4a W369F and *Cc*Cry4a W369F at different HDX-MS experimental conditions. All proteins were kept for 3 hr at 8 µM under the experimental conditions chosen for the HDX-MS experiments (LS, LSM, DS, DSM) with the exception of the samples which were kept on ice in the dark. The PageRuler Prestained Protein Ladder (Thermo Scientific) was used with the molecular weight markers in kDa as indicated.

4.2.2 Uptake comparison

Figures 4.4, 4.5, and 4.6 show the relative fractional uptake (RFU) (which is the relative uptake with respect to an undeuterated control sample calculated by dividing the deuterium level in Da by the total number of backbone amide hydrogens that could have become deuterated (which is equal to the number of amino acids, with the exception of proline residues and N-terminal residues) [6, 7] and here expressed as a percentage of each peptide's maximum possible uptake) and the RFU differences (ΔRFU) for several state comparisons for *Er*Cry4a WT, *Er*Cry4a W369F, and *Er*Cry4a R356F, respectively. The same figures for the the other four proteins investigated can be found in the Appendix, see Figures A.10-A.13. These plots have not yet been subjected to statistical data analysis which will follow in later sections

of this chapter. All plots, both the ones in the main text as well as those in the Appendix, show very similar RFU values in panel A of the three figures, suggesting similar folding of all proteins, including the apo-protein *Er*Cry4a R356E. The latter displays one region that seems to have a different uptake in comparison to all other proteins containing FAD. The so-called phosphate-binding loop (PBL, residues 231–248) [8] is already highly deuterated started at the earliest labelling time point. All proteins containing FAD display high uptake values and early saturation in the C-terminal region as well as multiple areas that experience protection upon light-exposure, indicated in blue in the relative fractional uptake differences plots, whereas there are barely any areas visible in which de-protection is observed.

In contrast to the experiments presented in chapter 3, a different set of labelling time points were chosen. All data sets have results for the 15 s, 1 min, 5 min, and 30 min labelling time points. The 1 hr labelling time point was not chosen for the present comparison as it did not yield much additional information. For *Er*Cry4a WT and W369F two shorter time points were added, 5 s and approximately '0.5' s. The latter labelling time point was reached by performing a 5 s labelling time point on ice rather than at room temperature. A reduction from 23°C to 4°C (on ice) to generate a ten-fold reduction in solvent-exchange rate for HDX-MS experiments was previously used elsewhere [9–11]. These short time points were chosen as a means to elucidate early structural changes in the C-terminal region. However, as Figures 4.4A and 4.5A show, the C-terminal tail was already saturated with high deuterium exchange at even these very early time points.

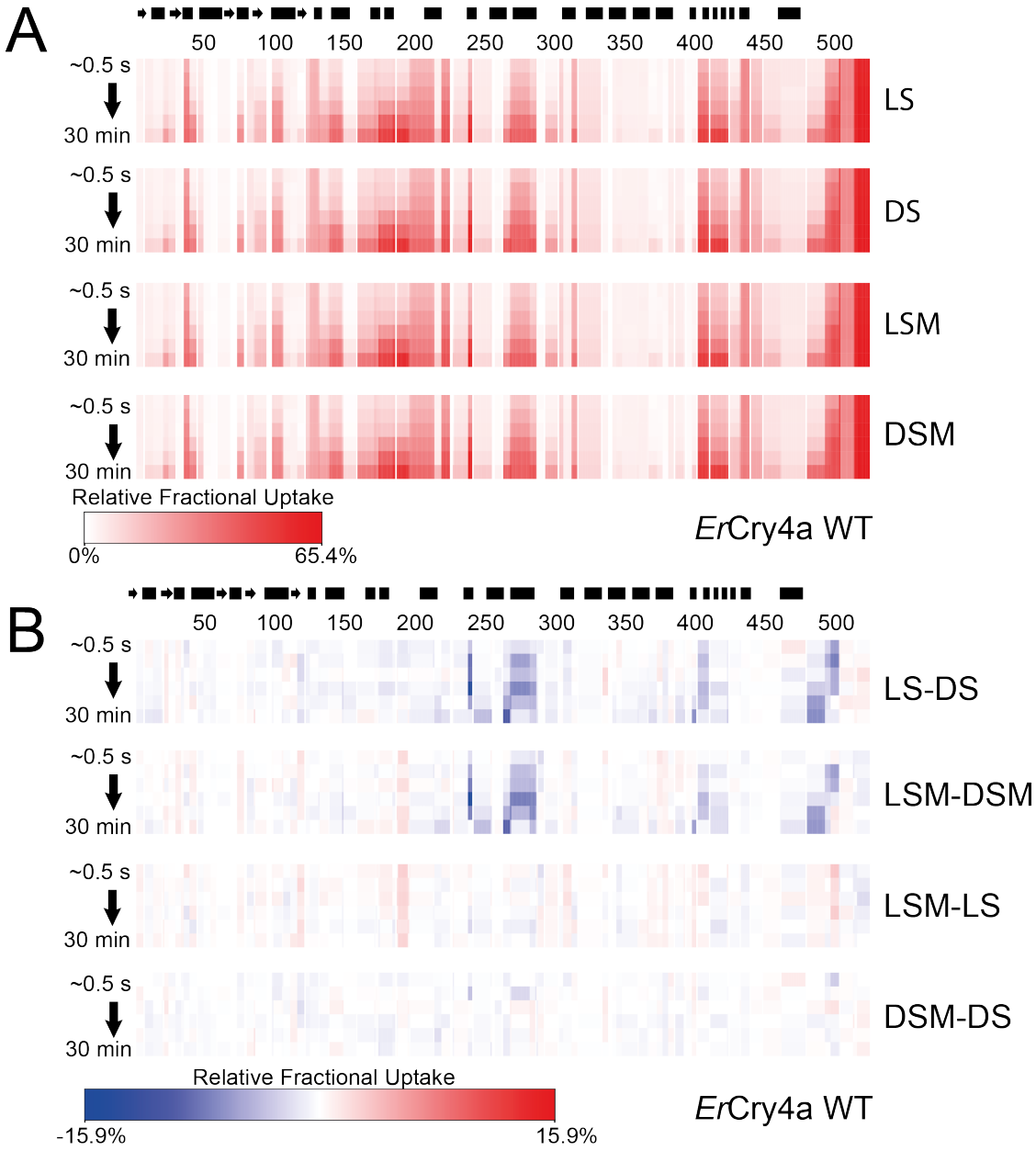

Figure 4.4: Relative fractional uptake of *Er*Cry4a WT. (**A**) Relative fractional uptake across the entire sequence measured at the peptide level. (**B**) Relative fractional uptake differences between states as indicated. As per usual convention, the N-terminus is situated on the left side and the C-terminus on the right. The time points displayed for both (A) and (B) are ∼0.5 s, 5 s, 15 s, 1 min, 5 min, and 30 min. The relative fractional uptake is calculated by dividing the deuterium level in Da by the total number of backbone amide hydrogens that could have become deuterated (which is equal to the number of amino acids, with the exception of proline residues and N-terminal residues) [6, 7]. For overlapping peptides, the data displayed is determined by the shortest peptide. Beta sheets (arrows) and alpha-helices (rectangles) are indicated above the sequences [12]. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

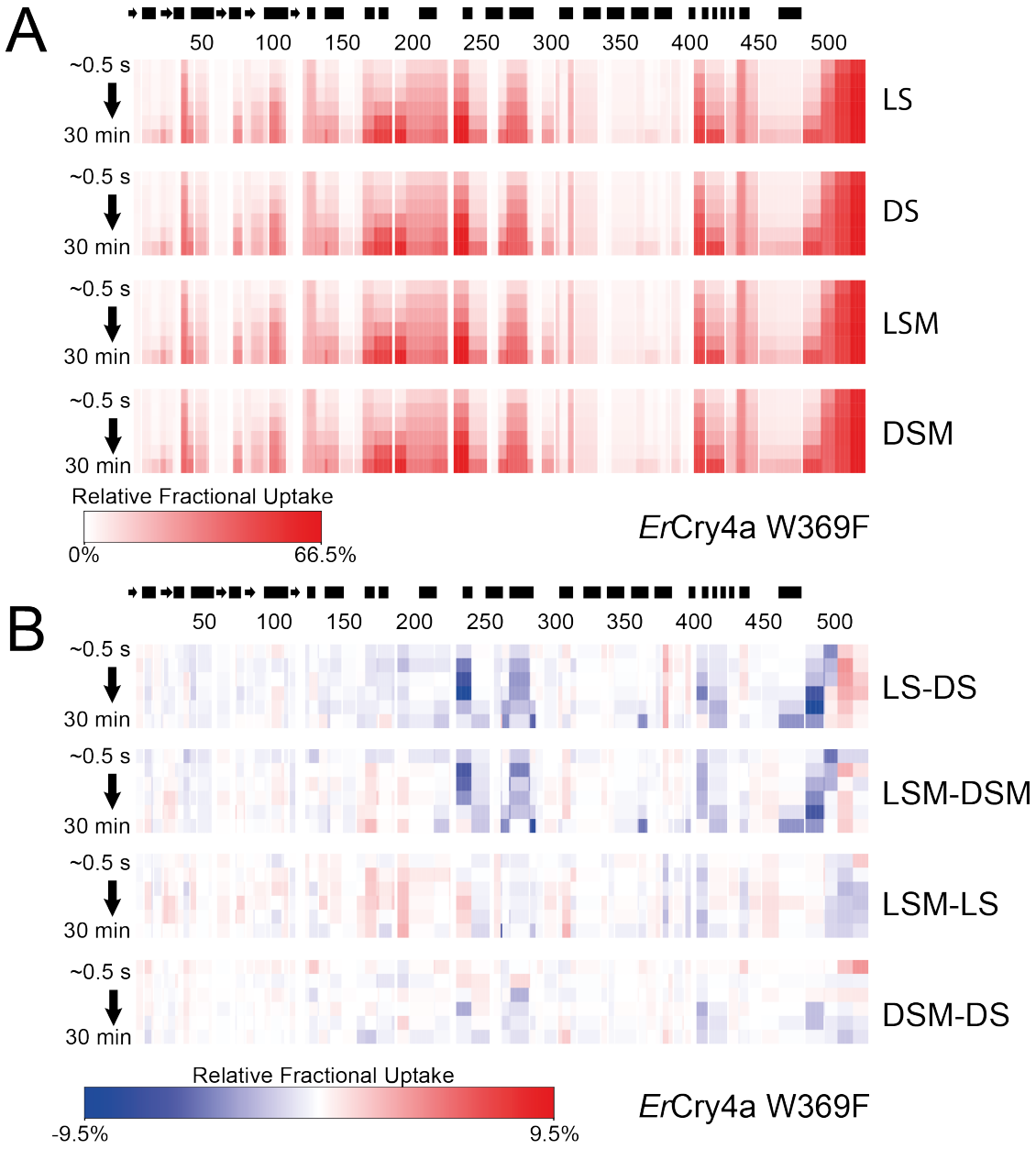

Figure 4.5: Relative fractional uptake of *Er*Cry4a W369F. (**A**) Relative fractional uptake across the entire sequence measured at the peptide level. (**B**) Relative fractional uptake differences between states as indicated. As per usual convention, the N-terminus is situated on the left side and the C-terminus on the right. The time points displayed for both (A) and (B) are ∼0.5 s, 5 s, 15 s, 1 min, 5 min, and 30 min. The relative fractional uptake is calculated by dividing the deuterium level in Da by the total number of backbone amide hydrogens that could have become deuterated (which is equal to the number of amino acids, with the exception of proline residues and N-terminal residues) [6, 7]. For overlapping peptides, the data displayed is determined by the shortest peptide. Beta sheets (arrows) and alpha-helices (rectangles) are indicated above the sequences [12]. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

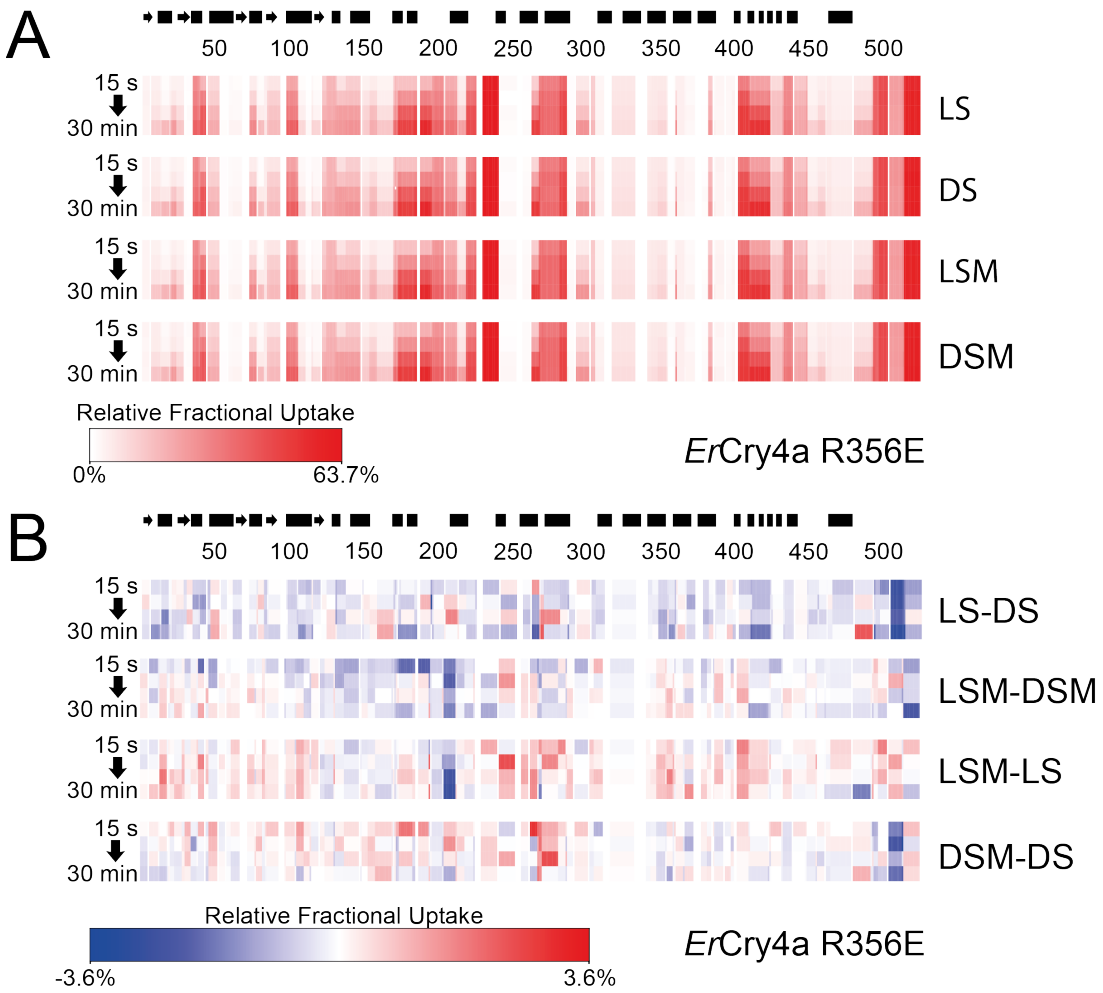

Figure 4.6: Relative fractional uptake of *Er*Cry4a R356E. (**A**) Relative fractional uptake across the entire sequence measured at the peptide level. (**B**) Relative fractional uptake differences between states as indicated. As per usual convention, the N-terminus is situated on the left side and the C-terminus on the right. The time points displayed for both (A) and (B) are ∼0.5 s, 5 s, 15 s, 1 min, 5 min, and 30 min. The relative fractional uptake is calculated by dividing the deuterium level in Da by the total number of backbone amide hydrogens that could have become deuterated (which is equal to the number of amino acids, with the exception of proline residues and N-terminal residues) [6, 7]. For overlapping peptides, the data displayed is determined by the shortest peptide. Beta sheets (arrows) and alpha-helices (rectangles) are indicated above the sequences [12]. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

4.2.3 Statistical data analysis

Deuteros

Further data processing using the Deuteros 2.0 software [5, 13] was performed. Deuteros has been specifically designed to be used post-DynamX for statistical analysis. It calculates confidence intervals and applies a Welch's two-tailed t-test. Additionally, it controls the number of false positives by correcting for multiple testing. However, all data analysis is performed separately for each time point, not taking the temporal component of HDX-MS data into account. For all the data discussed in the following which were analysed within Deuteros, we applied the hybrid significance test within Deuteros which first calculates if the deuterium uptake difference between two states reaches a user-defined significance level and subsequently applies Welch's t-test. Deuteros offers several visualisation tools, one of which is the so-called Woods plot as displayed in Figure 4.7. The calculated confidence interval is displayed as a dashed line and a p-value cutoff of <0.01 was applied. A rather low p-value cutoff was chosen to avoid over-interpretation into LSM-LS and DSM-DS differences as due to their rarity, the risk of interpreting too much into what might be false positives is increased. The peptides passing both the confidence interval and the two-tailed t-test as significant are displayed in blue [5].

Figure 4.7 displays one such Woods plot for *Er*Cry4a WT for the LS-DS difference. It shows the results of six different labelling times or exposures across the whole protein sequence. This figure serves as an example figure of what was a typical finding for all the data sets that will be discussed in detail in this chapter: mostly protection was observed upon exposure to light. This means that exposure to light resulted in a rearranged protein structure in which several parts of the protein were less surface accessible than when kept in the dark. Figure 4.7 also includes the two earliest time points, $\sim$0.5 s and 5 s which were only recorded for *Er*Cry4a WT and W369F. This figure exemplifies that these two additional time points, initially recorded to primarily learn more about the C-terminal tail, do not provide additional information aside from showing a few significant peptides in the same areas that show significant peptides for the later labelling time points.

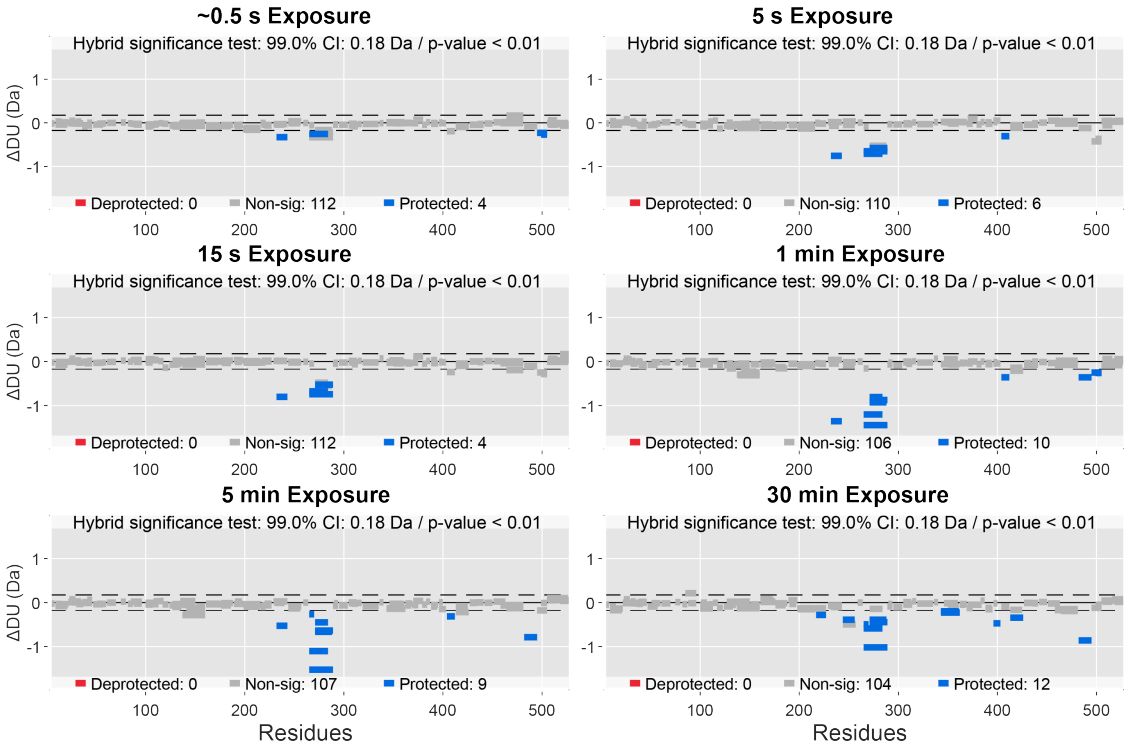

Figure 4.7: Deuteros Woods plot of *Er*Cry4a WT, LS-DS. Woods plots as the result of applying the hybrid significance test with a calculated confidence interval (CI) of 0.18 Da and a chosen p-value cutoff of 0.01. The plots include all peptides within the protein that were exported from DynamX after initial manual data analysis. All significant peptides display protection.

Using the Woods plots, Shah has created a range of RFU difference plots. In addition to the information the Woods plots provide, these plots indicate the amino acid region covered by each peptide, the relative uptake difference values divided by the maximum number of exchangeable hydrogens (ΔRFU), the named regions of the protein as well as alpha-helices and beta-sheets, making the differences in uptake more directly comparable. Figures A.22-A.25, A.27-A.30 , and A.32-A.35 show these plots separated by their labelling time points, directly mirroring the Woods plots, and Figures A.21, A.26 , and A.31 combine all time points in one plot by only displaying the peptides with the largest ΔRFU values across all time points. Unfortunately, due to the large size of these data sets, font sizes were sacrificed for a coherent display of all data for one bird cryptochrome species on one page. The figures have been used as the basis for the creation of Table 4.1 which is a collection of all significant peptides divided into regions of interest. This table is

mostly concerned with overall changes, not including the information at which time points which peptides have been found to be significant. In the following, each region that experiences changes leading to several overlapping significant peptides covering this region will be discussed separately. If not indicated otherwise, all discussed significant peptides in the following mark protection events where the significant peptide has become less solvent accessible upon exposure to either light or a magnet. The descriptions of low, medium, and high uptake are used as defined in the caption of the referenced table.

Table 4.1: Regions covered by significant peptides of six different cryptochrome 4a species after statistical data analysis in Deuteros. *Er*Cry4a R356E was not included in this table as no significant differences in relative fractional uptake differences (ΔRFU) were found. Unless otherwise indicated, only one peptide is covering the region. Low protection is defined as displaying ΔRFU between 0.02-0.04 Da, medium protection as between 0.04-0.08 Da, and high protection as above 0.08 Da. All numbers are given in Da. The table was jointly created by the author and Shah.

Region	***Er*Cry4a WT**	***Er*Cry4a W369F**	***Cc*Cry4a WT**	***Cc*Cry4a W369F**	***Gg*Cry4a WT**	***Gg*Cry4a W369F**
α5 - α6 (region 136-161)	Low protection, largely α6 (148-160), [LSM-DSM]	-	Low protection, loop (143-154), [LS-DS, LSM-DSM]	Low protection, loop (143-154), [LS-DS, LSM-DSM]	Low protection, loop (143-154), [LS-DS, LSM-DSM]	-
α8, phosphate-binding loop, α9 (region 217-256)	High protection (>0.14 Da), surrounding area also shows protection, (217-256), [LS-DS, LSM-DSM]	High protection, (232-243), [LS-DS], no significant peptides in [LSM-DSM]	High protection, (232-243), [LS-DS, LSM-DSM]	High protection, 2 peptides (232-243, 229-243), [LS-DS, LSM - DSM]	High protection (>0.14 Da), (232-243), [LS-DS, LSM-DSM]	High protection (>0.14 Da), 2 peptides (229-243, 232-243), [LS-DS, LSM-DSM], & low to medium protection, 2 peptides (244-255/256), [LS-DS, LSM - DSM]

Continued on next page

Table 4.1 – continued from previous page

Region	***Er*Cry4a WT**	***Er*Cry4a W369F**	***Cc*Cry4a WT**	***Cc*Cry4a W369F**	***Gg*Cry4a WT**	***Gg*Cry4a W369F**
α10, protrusion motif, α11 (region 264-289)	Medium protection, 5 overlapping peptides (265-289), [LS-DS, LSM-DSM] & high protection, additional peptide covering α10 (264-270), [LS-DS, LSM-DSM]	High protection, (285-289), [LSM-DSM] & low to medium protection, 5 peptides (265-289), [LS-DS], 4 peptides (254-289), [LSM-DSM]	Low to medium protection, (265-289), [LS-DS (3 peptides), LSM-DSM]	Medium protection, (264-289), [LS-DS (4 peptides), [LSM-DSM (7 peptides)]	Medium to high protection (variable range of values), 7 peptides (region 264-289), [LS-DS, LSM-DSM] & low protection, (271-289), [DSM-DS]	High protection (>0.14 Da), (265-272), [LS-DS, LSM-DSM] & medium to high protection (variable range of values), 10 peptides (region 264-289), [LS-DS, LSM-DSM]
Loop (region 305-314)	-	-	-	-	-	Low protection, (305-314), [LS-DS, LSM-DSM]
α14 (region 362-369)	-	Low to medium protection, 3 peptides (362-369), [LS-DS (only low protection), LSM-DSM]	-	-	-	Low to medium protection, 2 peptides (362-368/9), [LS-DS, LSM-DSM]

Continued on next page

Table 4.1 – continued from previous page

Region	***Er*Cry4a WT**	***Er*Cry4a W369F**	***Cc*Cry4a WT**	***Cc*Cry4a W369F**	***Gg*Cry4a WT**	***Gg*Cry4a W369F**
384-403 (includes α16)	Medium protection, (396-403), [LS-DS, LSM-DSM]	De-protection, loop (384-388), [LS-DS]	-	-	-	Medium protection, loop (384-388), [LS-DS, LSM-DSM] & medium to high protection, α16 (396-400), [LS-DS (high protection), LSM-DSM]
404-426	Low to medium protection, 2 peptides (404-412, 413-426), [LS-DS, LSM-DSM]	Low to medium protection, (404-412, 413-426), [LS-DS (2 peptides), LSM-DSM]	Low protection, (404-412), [LS-DS, LSM-DSM]	Medium protection, (404-412), [LS-DS, LSM-DSM] & low protection, (413-425), [LSM-LS]	Low protection, (401-412), [LSM-DSM]	Low to medium protection, (404-426), [LS-DS (5 peptides), LSM-DSM (3 peptides)]

Continued on next page

Table 4.1 – continued from previous page

Region	***Er*Cry4a WT**	***Er*Cry4a W369F**	***Cc*Cry4a WT**	***Cc*Cry4a W369F**	***Gg*Cry4a WT**	***Gg*Cry4a W369F**
α22 (region 458-495)	Medium protection, (482-495), [LS-DS, LSM-DSM]	High protection (482-495), [LS-DS] & low protection, (458/459-481, 482-495), [LSM-DSM]	Medium protection, (490-495), [LS-DS]	Medium to high protection (and a singular peptide of low protection), (region 459-495), [LS-DS (5 peptides), LSM-DSM (3 peptides)] & high protection (>0.14 Da), (490-495), [LS-DS, LSM-DSM]	Medium protection, (482-495), [LS-DS, LSM-DSM]	Medium to high protection, 3 peptides (2 of which display >0.12 Da) (464-481, 471-481, 482-495), [LS-DS, LSM-DSM]
C-terminal loop (496-505)	Low to medium protection, 2 peptides (495-505, 499-505), [LS-DS]	Low to medium protection, 2 peptides (495-506, 496-505), [LS-DS]	Low protection, (496-505/506), [LS-DS, LSM-DSM (2 peptides)]	Low protection, (495/6-505/6), [LS-DS (4 peptides), LSM-DSM]	Variable degrees of protection, 2 peptides (496-505/506), protection is higher in [LS-DS] than [LSM-DSM]	Low protection, 4 overlapping peptides (region 495-506), [LSM-DSM]

Continued on next page

Table 4.1 – continued from previous page

Region	***Er*Cry4a WT**	***Er*Cry4a W369F**	***Cc*Cry4a WT**	***Cc*Cry4a W369F**	***Gg*Cry4a WT**	***Gg*Cry4a W369F**
C-terminal (α23 - end)	-	De-protection, 2 peptides (505-526, 506-527), [LS-DS], & low protection, (505-526), [LSM-LS]	-	Low protection, 2 peptides (506-527, 507-529), [LSM-LS]	-	-
Other peptides	-	De-protection, (190-198), [LSM-LS]	De-protection, (129-136), [LSM-DSM]	Low protection, (161-173), [DSM-DS]	Low protection, (36-46), [LSM-DSM]	Medium protection, (75-81), [LS-DS] & low protection, (177-186), [DSM-DS]

Region 136-161: This region is at least partially covered for all three WT proteins, as well as for *Cc*Cry4a W369F for the LS-DS differences and the LSM-DSM differences with the exception of *Er*Cry4a WT, where it does not appear for the LS-DS comparison. This region mostly concerns a loop between α5 and α6. Only *Er*Cry4a WT is an exception where the significant peptide covers mostly α6 rather than the loop.

Phosphate-binding loop (231–248): The phosphate-binding loop (PBL) spans residues 231-248 for *Er*Cry4a WT [8] and is assumed to cover a very similar amino acid region in the other bird cryptochromes. The coverage of this region is similar within the chicken and quail proteins as well as in *Er*Cry4a WT: one or two peptides of high ΔRFU values are found in this protein region for both the LS-DS and LSM-DSM differences. Only for *Er*Cry4a W369F there is no significant peptide present for the LSM-DSM comparison. In *Er*Cry4a WT the area covered with significant peptides extents to α8 whereas in *Gg*Cry4a W369F the same is the case for the other side of the loop where peptides extent to cover α9 as well.

Protrusion motif (273–282): The protrusion loop or more often called the protrusion motif (PM) spans residues 273-282 for *Er*Cry4a WT [8] and is assumed to cover a very similar amino acid region in the other bird cryptochromes. The coverage of this region stands out as it is covered by a lot of overlapping and comparatively long peptides for the LS-DS and LSM-DSM comparisons. Only *Er*Cry4a W369F and *Cc*Cry4a WT display fewer peptides covering the area. However, this is not a direct reflection of the extent of protection in this area as the total number of peptides found for this area, significant or not, varies for different proteins. *Er*Cry4a WT as well as *Gg*Cry4a WT and W369F have comparatively higher ΔRFU values for the peptides covering the area, with the highest values ($<$0.14 Da) found in *Gg*Cry4a W369F. These values are among the highest found overall across the entire sequences of all proteins. Additionally, *Gg*Cry4a WT has one significant peptide coming up covering this region for the DSM-DS comparison.

Region 305-314: This region is only partially covered within *Gg*Cry4a W369F and only for the longest labelling time point of 30 min.

α14 ($\sim$353-363): Only significant in two proteins, *Er*Cry4a W369F and *Gg*Cry4a W369F. However, the coverage in this region only occurs after 30 min of labelling time. The reason it is not visible in *Cc*Cry4a W369F as well could be that it takes longer for the protein region to be labelled in this protein in the LS. Additionally, this protein area is slightly more buried within the protein and would therefore be expected to become deuterated more slowly.

α16 and surrounding regions (384-403): Only *Er*Cry4a and *Gg*Cry4a display significant peptides in this area. *Er*Cry4a W369F stands out by displaying a deprotection event (gaining more surface accessibility) in region 384-388 for the LS-DS comparison where *Gg*Cry4a W369F only displays the usual protection (less surface accessibility) in the same region for both the LDS-DS and LSM-DSM comparisons. α16 is only protected in *Er*Cry4a WT and *Gg*Cry4a W369F. The changes in this region might happen due to the changes observed in the neighbouring loop region 404-415 and adjacent α17 as they are only observed if at least two overlapping peptides in the latter region have been found to be significant (see below). Again, this region is more buried within the protein, similar to α14.

α17 and both surrounding loops ($\sim$404-426): This region experiences significant changes in all proteins, usually in one or two peptides for the LS-DS and LSM-DSM comparisons (with the exception of *Gg*Cry4a WT with no peptides for the LS-DS comparison and *Gg*Cry4a W369F where there are several peptides in both comparisons). *Gg*Cry4a displays only one significant peptide for each of the two comparisons in both proteins.

α22 (471-495): This region is generally covered in all six proteins. All WT proteins display one significant peptide covering the region for all LS-DS and LSM-DSM comparisons with only *Cc*Cry4a WT having no significant peptide in

the LSM-DSM comparison. In the W369F proteins, the experienced changes are larger and/or more peptides overlap in this region in comparison to their WT proteins. Especially *Cc*Cry4a and *Gg*Cry4a W369F show large ΔRFU differences with the largest in 482-495.

Loop 496-505: The loop between α22 and α23 experiences changes in all proteins. The only comparisons which do not lead to significant peptides are the LSM-DSM comparison for *Er*Cry4a both WT and W369F and the LS-DS comparison for *Gg*Cry4a W369F. The observed changes could be due to the changes observed in the neighbouring α22 which for some proteins experienced very large ΔRFU values.

α23-end: In this region *Er*Cry4a W369F experiences de-protection for the LS-DS comparison and protection for the LSM-LS comparison. Keeping in mind that this region of the protein is already saturated (fully deuterated) at early time points, the findings in this region, although of interest, do not carry the same meaning as findings for less easily saturated areas of the protein. However, these findings would suggest that the addition of the magnet reduces the C-terminal accessibility which is otherwise generally increased by light-induced changes. *Cc*Cry4a W369F displays the same protection for the LSM-LS comparison whereas no such changes occur in *Gg*Cry4a.

Comparison of WT and W369F: For *Er*Cry4a, there are a few clear differences between WT and W369F. The W369F protein experiences protection events near the mutation site (however, they are only visible after 30 min of labelling time) and de-protection in the C-terminal region for the first two labelling time points, 15 s and 1 min. There are also LSM-LS differences for the W369F mutant: significant peptides are observed for 505-526 (LSM-LS, only at 15 s) and for 190-198 (LSM-LS, only at 5 min). None of these changes were observed in the WT protein.

For the W369F protein, the 1 min and 5 min labelling time points show almost no significant peptides for the LSM-DSM comparison as opposed to significant peptides that have been found for the LS-DS comparison. Overall, the WT protein

shows less variability between the LSM-DSM and the LS-DS comparisons than the W369F protein.

For the second migratory bird protein, *Cc*Cry4a, there are no differences at the mutation site and no protection at the C-terminus, but for *Cc*Cry4a W369F there are significant peptides for 506-526 and 507-529 for the LSM-LS comparison at the 5 min labelling time point. Two further peptides which were found as a result of the addition of a magnet were also only found in the W369F mutant. Similar to *Er*Cry4a, the W369F protein experiences more variability in terms of which changes are visible at which labelling time point and at which ΔRFU values. Especially interesting is the area between residues 459-489 which experiences protection in many overlapping peptides in the W369F mutant, but not in the WT.

For the non-migratory bird protein, *Gg*Cry4a, there are changes at the mutation site for *Gg*Cry4a W369F. The region experiences protection for the LS-DS and LSM-DSM comparisons at the 30 min labelling time point, similar to what was found for *Er*Cry4a W369F. However, neither de-protection at the C-terminus nor any significant peptides for the LSM-LS comparison were observed. Only one peptides of significance for the DSM-DS comparison was found for *Gg*Cry4a W369F (177-186). Again, as in the other two bird proteins, a higher degree of variability of the results is observed for the W369F mutant in comparison to the WT, especially in the region covering α14-α17. Interestingly, there are no significant peptides at all for the LSM-DSM comparison for the 15 s labelling time point of both *Gg*Cry4a WT and W369F despite the LS-DS comparison displaying clear regions where differences are observed.

Summing up: Several areas covered by significant peptides for the LS-DS comparison have been listed and analysed in detail. However, some areas are of heightened interest. The PM is one such area, being consistently covered by several, overlapping peptides displaying protection, similar to the nearby PBL. It is conceivable, also given the length of the PBL, that it might undergo a movement towards the PM and surrounding α-helices which would end up covering all of

these areas to a certain extent which could explain the amount of protection observed. Similarly, the region spanning α16, α17, and their connecting loop regularly exhibits significance in all proteins in many of the labelling time points (especially pronounced for *Gg*Cry4a W369F) and could potentially be covered as a result of movements within the C-terminal tail as both of these regions (mostly) display protection events. Interesting to mention in this context is that Schuhmann *et al.* recently performed molecular dynamics simulations on pigeon cryptochrome 4 finding that the PBL might act as a gate for access to the cofactor FAD [14]. However, these interpretations are only attempts to explain the data presented here and require further experimental studies. Another interesting result is that the *Gg*Cry4a W369F mutant experiences such large protection changes in the C-terminal region, especially in α22.

There are clear differences between WT and W369F proteins with the latter generally displaying a greater variability in terms of numbers of overlapping peptides, ΔRFU values, and consistency of significant protein areas across labelling time points. They also span a greater part of the protein sequence in *Er*Cry4a and *Gg*Cry4a in comparison to *Cc*Cry4a. *Cc*Cry4a WT and *Er*Cry4a W369F stand out by being the two proteins that experience any de-protection in a few peptides while all other observed significant peptides exclusively depict protection. Further discussion, including of the light-induced differences, can be found below where the results of an alternative analysis method are presented.

HD-eXplosion

As the data sets presented here contain an enormous amount of information, we have applied different statistical analysis approaches in order to make the data sets more accessible and facilitate the data analysis. Before using Deuteros which led to the analysis and results described in detail above, we also employed the online tool HD-eXplosion [15] which uses only a Welch's t-test to evaluate the statistical significance of our data. Since this is a slightly less rigorous method than used in Deuteros as it does not correct for multiple testing which Deuteros does (although

the perfect method for this kind of analysis still seems to be under debate) [5], the results using HD-eXplosion will only be briefly discussed here. The reason this approach can be helpful in assisting the data analysis is that it leads to a different type of display, the so-called chiclet plots, as displayed in Figures A.14-A.20 in the Appendix with the differences between LS and DS for *Er*Cry4a WT, *Er*Cry4a W369F, and *Er*Cry4a R356E displayed in Figure 4.8 as a slightly truncated version of the original figures. It is important to keep in mind that the same peptides or protein regions are not exactly vertically above each other in the figure as the three protein data sets have different numbers of total peptides found.

Generally, these plots display every single peptide found within the data sets on one axis and the different labelling time points on the other axis, making the time evolution of the significance of certain peptides more accessible. However, the widths of the regions of change correspond to the number of peptides found in each region and are not a measure of the significance of the observed changes within that area. Another problem of this display is the size of the plot as it becomes rather large for large proteins. However, the visual examination of the evolution of significance at different labelling time points is of value for the data interpretation and helped with the Deuteros analysis above. Figure 4.8 nicely shows the de-protection observed for the W369F mutant at earlier labelling time points which was not found in the WT. It also visually highlights how the PBL and PM experience protection at almost every time point, which could be interpreted as a resulting stable structure after structural rearrangement with a few less exposed areas which lead to the respective areas in the LS getting increasingly deuterated with every time point in comparison to the same areas in the DS which are significantly more exposed starting already from early time points. Furthermore, the R356E mutant displays only two significant peptides which seem to be rather randomly distributed across time points and the sequence pointing towards them being outliers that passed the significance test. Since HD-eXplosion uses a less elaborate method of significance testing than Deuteros, the results vary in some protein areas. One example is *Cc*Cry4a W369F displaying some de-protection for the LS-DS comparison which is

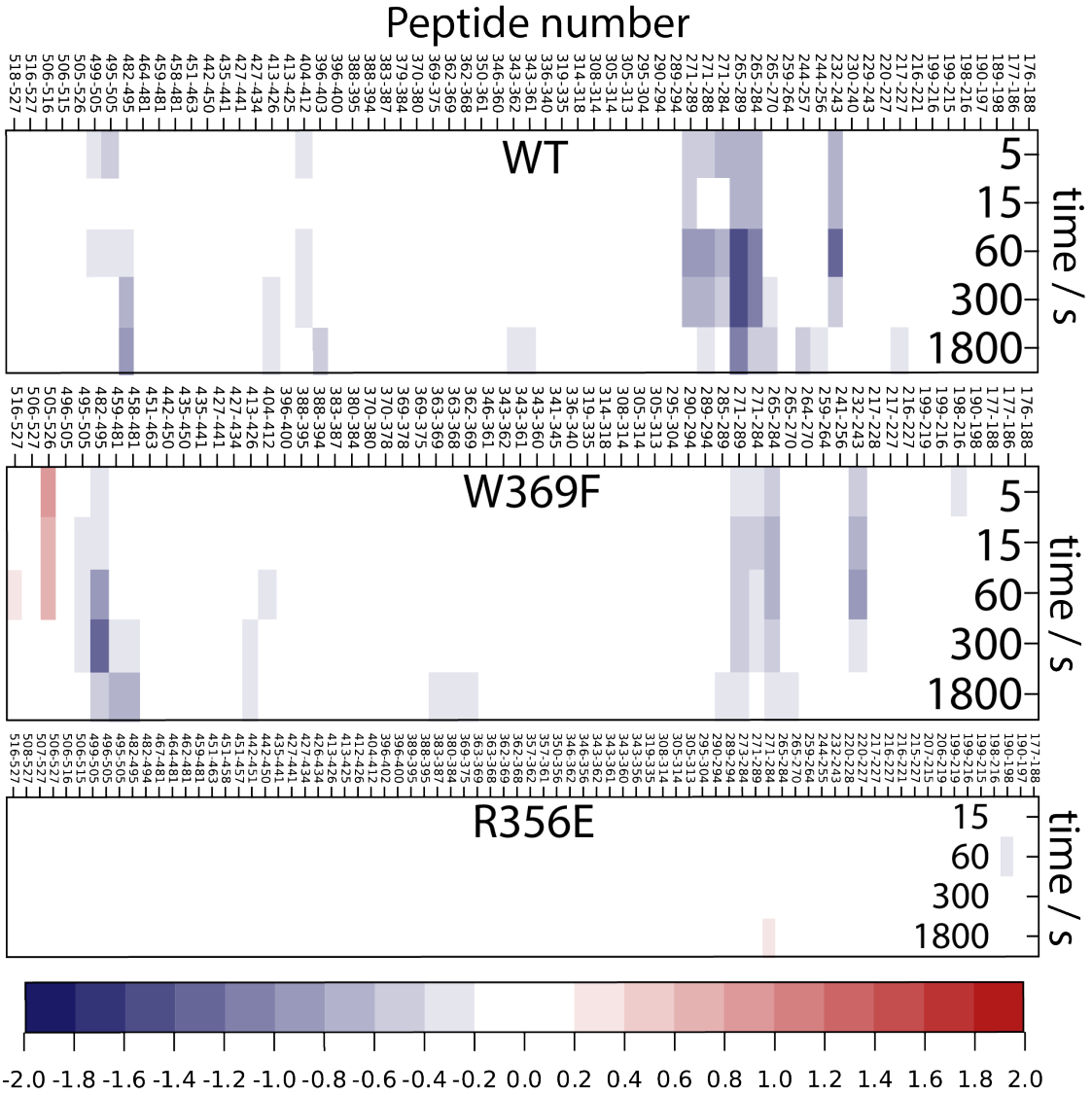

Figure 4.8: Peptides of significant uptake of *Er*Cry4a WT, W369F, and R356E for the LS-DS comparison. HD-eXplosion chiclet plots [15]. Significance filtering was applied by employing a cutoff p-value of 0.01. The N-terminal end of the sequence is at the right side of the plot and the C-terminal end on the left corresponding to the order of peptides in Figures A.14, A.15, and A.16 in the Appendix from which these plots are excerpts.

not the case in the Deuteros results. Another difference is that differences between states with and without magnets come up as significant more often using this method. However, since the method of significance testing is less elaborate, the focus of this chapter will be on the Deuteros and hdxtstats (see below) analysis.

Probing magnetic field-induced impact on conformational changes - introducing hdxstats

One important reason to use different statistical approaches to analyse our findings, as pointed out in chapter 3, is the need to analyse the LSM-LS and DSM-DS differences more rigorously. Therefore, a third method of statistical data analysis was employed in collaboration with Oliver Crook, a Florence Nightingale Bicentennial Fellow in the Department of Statistics of the University of Oxford, who recently designed a statistical analysis tool specifically for HDX-MS data called hdxstats [16]. The advantage of this new software, which distinguishes it from other HDX-MS data analysis tools, is that it takes the temporal component of HDX-MS data into account using an empirical Bayes methodology while also providing improved statistical power which reduces the number of false positives.

A comparison of the statistical analysis performed using Deuteros and hdxstats on the same seven data sets is displayed in tables 4.2, 4.3, and 4.4 for the *Er*Cry4a, *Cc*Cry4a, and *Gg*Cry4a proteins, respectively. The tables combine a list of all peptides that were found to be significant either in Deuteros or in HDX-stats in all seven proteins. Where there were no significant peptides found, the columns for the particular state comparisons are not displayed.

One of the first observations is that each table has a different length or number of listed peptides with the table containing the *Gg*Cry4a results being the longest. This is due to the pepsin being rather unspecific when digesting the proteins. However, it is interesting that one protein seems to yield a lot more significant peptides than the other two. Whereas for the robin and the quail cryptochromes the peptides found to be significant by both Deuteros and hdxstats seem to roughly line up, this is not the case for the chicken cryptochrome where hdxstats elongates the list by adding more overlapping peptides in significant regions that were not found in Deuteros. Another

finding when comparing all three tables is that the first peptides up until residues 129 (four peptides), 81 (four peptides), and 36 (six peptides) for *Er*Cry4a, *Cc*Cry4a, and *Gg*Cry4a, respectively, have only been found to be significant in hdxstats.

Table 4.2: Significant peptides of *Er*Cry4a WT and W369F. Significant peptides found by hdxstats are denoted as 'x'. Deuteros (p-value <0.01) and hdxstats results (empirical Bayes false discovery rate cutoff value <0.01; 0.01<value<0.05 marked with an asterisk) are compared. The Deuteros results list all time points in which the peptides were significant. If peptides were present in several charge states, the hdxstats results classify the result as <0.01 if this was true for at least one charge state. The table does not differentiate between charge states as some charge states were only present for some peptides (depending on the initial raw data). If there were no significant peptides found for any of the state differences, the respective columns are not displayed. The results for *Er*Cry4a R356E are not listed in the table. The few peptides of significance are listed in the following with their corresponding state comparisons: 1. Deuteros: no signifcant peptides. 2. hdxstats: [LS-DS]: 38-46 x*, 271-284 x; [LSM-DSM]: none; [LSM-LS]: 38-46 x*; [DSM-DS]: 265-284 x.

	***Er*Cry4a WT**				***Er*Cry4a W369F**						
Peptide	**Deuteros**		**hdxstats**		**Deuteros**			**hdxstats**			
	LS-DS	LSM-DSM	LS-DS	LSM-DSM	LS-DS	LSM-DSM	LSM-LS	LS-DS	LSM-DSM	LSM-LS	DSM-DS
61-70									x*	x*	
74-81										x*	
75-81										x*	
113-117										x	
129-142					1 min, 30 min			x			
138-160			x*								
141-160		5 min	x*	x*							
148-160		1 min, 30 min	x*	x*							

Continued on next page

Table 4.2 – continued from previous page

	***Er*Cry4a WT**				***Er*Cry4a W369F**						
Peptide	**Deuteros**		**hdxstats**		**Deuteros**			**hdxstats**			
	LS-DS	LSM-DSM	LS-DS	LSM-DSM	LS-DS	LSM-DSM	LSM-LS	LS-DS	LSM-DSM	LSM-LS	DSM-DS
151-160			x	x*				x			
167-175							1 min				
176-188					15 s, 1 min						
190-198							5 min, 30 min				
198-216					30 min			x			
199-219					30 min						
216-221			x*	x*							
216-227					30 min			x	x*		
217-227	30 min	30 min		x*	30 min			x	x		
220-227		30 min									
232-243	15 s, 1 min, 5 min	15 s, 1 min, 5 min	x	x	15 s, 1 min, 5 min			x	x		
241-256								x			
244-256	30 min	30 min	x	x							
244-257			x								
264-270						30 min			x		

Continued on next page

Table 4.2 – continued from previous page

	***Er*Cry4a WT**				***Er*Cry4a W369F**						
Peptide	**Deuteros**		**hdxstats**		**Deuteros**			**hdxstats**			
	LS-DS	LSM-DSM	LS-DS	LSM-DSM	LS-DS	LSM-DSM	LSM-LS	LS-DS	LSM-DSM	LSM-LS	DSM-DS
265-270	15 s, 1 min, 5 min, 30 min	5 min, 30 min	x	x	5 min, 30 min	30 min		x	x		
265-284	15 s, 1 min, 5 min, 30 min	1 min, 5 min, 30 min	x	x	15 s, 1 min, 5 min, 30 min	15 s, 5 min	5 min	x	x		
265-289	15 s, 1 min, 5 min, 30 min	15 s, 1 min, 5 min, 30 min	x	x							
271-284	1 min, 5 min	15 s, 1 min, 5 min	x	x	1 min, 5 min			x	x		
271-288	1 min, 5 min, 30 min	15 s, 1 min, 5 min, 30 min	x	x							
271-289	15 s, 1 min, 5 min, 30 min	15 s, 1 min, 5 min, 30 min	x	x	1 min, 5 min, 30 min	15 s, 30 min		x	x		
285-289					30 min	30 min		x	x		
308-314									x*		x
336-340			x								
343-361	30 min		x	x	5 min, 30 min						
343-362	30 min	30 min	x	x				x*	x*		

Continued on next page

Table 4.2 – continued from previous page

	***Er*Cry4a WT**				***Er*Cry4a W369F**						
Peptide	**Deuteros**		**hdxstats**		**Deuteros**			**hdxstats**			
	LS-DS	LSM-DSM	LS-DS	LSM-DSM	LS-DS	LSM-DSM	LSM-LS	LS-DS	LSM-DSM	LSM-LS	DSM-DS
346-361								x	x*		
362-369					30 min	30 min		x	x		
363-368					30 min	30 min		x	x	x*	
363-369					30 min	30 min		x	x		
369-375			x								
370-380				x							
380-384					30 min			x		x*	x
388-394			x	x							
388-395			x	x					x*		
396-400								x	x		
396-403	30 min	30 min	x	x							
404-412	1 min, 5 min	1 min	x	x	1 min	15 s, 5 min		x	x		x*
413-426	30 min	5 min			1 min, 5 min, 30 min			x	x*		
451-463							5 min				
458-481					5 min	5 min, 30 min		x	x		
459-481					5 min	30 min		x	x		

Continued on next page

Table 4.2 – continued from previous page

	***Er*Cry4a WT**				***Er*Cry4a W369F**						
Peptide	**Deuteros**		**hdxstats**		**Deuteros**			**hdxstats**			
	LS-DS	LSM-DSM	LS-DS	LSM-DSM	LS-DS	LSM-DSM	LSM-LS	LS-DS	LSM-DSM	LSM-LS	DSM-DS
482-495	1 min, 5 min, 30 min	1 min, 5 min, 30 min	x	x	1 min, 5 min	15 s, 30 min		x	x		
495-505	1 min		x	x	15 s			x	x		
496-505					15 s			x	x		
499-505	1 min		x	x							
505-526					15 s		15 s	x		x	
506-527					1 min			x		x	
516-527										x*	

For *Er*Cry4a WT, the magnet did not have a measurable influence as both Deuteros and hdxstats did not find any significant peptides. For *Cc*Cry4a WT, two peptides were found in Deuteros for LSM-LS (81-96, 515-529) and two for DSM-DS (451-463, 506-525). hdxstats found four peptides for LSM-LS (8-14, 9-21, 9-22, 88-9) and five peptides for DSM-DS (9-21, 9-22, 41-46, 137-142). However, all peptides found by Deuteros are only significant for one time point and all results found by hdxstats are only below a cutoff value of 0.05, but not below 0.01 with the exception of 137-142 for the DSM-DS difference. However, this peptide is assumed to be an outlier as it does not show up for any of the other state comparisons. It is interesting to note that most of these peptides can be found towards the N-terminus and two towards the C-terminal tail. For *Gg*Cry4a WT, Deuteros does not find any significant peptides for the LSM-LS comparison where hdxstats finds 10 peptides which is quite a large difference. These were found in the peptides 23-29, 113-123, 124-128, 265-270 (with a false discovery rate (FDR) below 0.01), 343-362, 346-362,

363-368 (FDR <0.01), 369-375, 383-378, and 388-395 (FDR <0.01). Three of these peptides have a FDR <0.01, one of which is in the protrusion motif region and the other two in regions that do not contain any significant peptides in the Deuteros analysis, in either of the state comparisons. However, these peptides are rather wildly scattered across the whole protein instead of covering certain regions. For the DSM-DS comparison, hdxstats found three significant peptides (124-128, 363-368, 388-395). One of the peptides (363-368) has a FDR <0.01 and has also been found in the LSM-LS comparison, but not in the LS-DS and LSM-DSM comparisons, suggesting this change could occur as the result of adding a magnet.

Analysing the LSM-LS and DSM-DS state comparisons for the W369F mutants of all three proteins, we find significant peptides for both comparisons in *Er*Cry4a W369F. For LSM-LS, Deuteros finds five peptides (167-175, 190-198, 265-284, 451-463, 505-526) with one of them coming up as significant for two time points (190-198) and hdxstats finds nine peptides (61-70, 74-81, 75-81, 113-117, 363-368, 380-384, 505-526, 506-527, 516-527). Interestingly, two of these peptides have a FDR <0.01 which both cover the C-terminal tail, supported by a third peptide with an FDR <0.05 in the same region. This also corresponds to one peptide that was found to be significant within the same region by Deuteros. For DSM-DS, Deuteros finds no significant peptides, but hdxstats finds three (308-314, 380-384, 404-412) with the first two being below a FDR of 0.01. Interestingly, peptide 380-384 comes up as significant for both light and dark states when a magnet was added.

For *Cc*Cry4a W369F, LSM-LS differences are found by Deuteros for 413-425, 506-522, and 507-529, each of which was only found for one time point where hdxstats finds no significant peptides. It is important to note here, that the centroids of the charge state distributions in the raw data for 506-522 and 507-529, based on which Deuteros calculates the uptake differences, were not completely representative of the actual distribution of these particular peptides. Therefore, resulting significant peptides need to be discussed with caution. For the DSM-DS differences, Deuteros finds only one peptide (161-173) and hdxstats finds the same peptide as well as 435-441. Several of these peptides seem to stand alone, neither

supported by surrounding significant peptides nor, if found by one analysis tool, by a significant peptide in the same position found by the other analysis tool. However, interesting regions are the 161-173 as this was recorded as a magnet-induced change in both analysis tools for the DSM-DS difference and the C-terminal region which was recorded for the LSM-LS difference in Deuteros.

Table 4.3: Significant peptides of *Cc*Cry4a WT and W369F. Significant peptides found by hdxstats are denoted as 'x'. Deuteros (p-value <0.01) and hdxstats results (empirical Bayes false discovery rate cutoff value <0.01; 0.01<value<0.05 marked with an asterisk) are compared. The Deuteros results list all time points in which the peptides were significant (listed in minutes unless otherwise indicated). If peptides were present in several charge states, the hdxstats results classify the result as <0.01 if this was true for at least one charge state. The table does not differentiate between charge states as some charge states were only present for some peptides (depending on the initial raw data). If there were no significant peptides found for any of the state differences, the respective columns are not displayed.

	***Cc*Cry4a WT**								***Cc*Cry4a W369F**						
Peptide	**Deuteros**				**hdxstats**				**Deuteros**				**hdxstats**		
	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	DSM-DS
8-14						x*	x*								
9-21							x*	x*							
9-22							x*	x*							
41-46								x*						x	
81-96			30												
88-96							x*								
100-107					x*										
100-109					x*										
100-112					x										

Continued on next page

Table 4.3 – continued from previous page

	***Cc*Cry4a WT**								***Cc*Cry4a W369F**						
Peptide	**Deuteros**				**hdxstats**				**Deuteros**				**hdxstats**		
	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	DSM-DS
129-136		5, 30				x									
137-142								x							
143-154	5	1			x	x			30	1			x	x*	
155-160					x	x							x	x	
159-175		15 s													
161-173												30			x
229-243									15 s, 1	1, 5			x	x	
232-243	1	1			x	x			15 s, 1	1, 5			x	x	
244-256														x*	
264-270					x	x			30	30			x	x	
265-270					x	x			30	30			x	x	
265-284	1	1			x	x			15 s, 1, 30	1, 5, 30			x	x	
271-284	1, 5				x	x			15 s, 1	1, 5			x	x	
271-289									1, 30	1, 5			x	x	
272-284									15 s, 1	1, 5			x	x	
273-284										15 s, 1			x	x	

Continued on next page

Table 4.3 – continued from previous page

	***Cc*Cry4a WT**								***Cc*Cry4a W369F**						
Peptide	**Deuteros**				**hdxstats**				**Deuteros**				**hdxstats**		
	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	DSM-DS
273-289	1				x	x									
285-289													x	x	
290-294								x*							
295-304					x*										
343-362						x									
369-375					x										
396-400					x									x*	
404-412	1, 5	1			x	x			1	1, 5			x	x	
413-425											5		x*		
435-441													x		x
451-463				5											
459-481									30	5			x	x	
464-481									5, 30				x	x	
471-481									5, 30	5			x	x	
482-489						x			1, 5, 30	5			x	x	
482-495					x	x									

Continued on next page

Table 4.3 – continued from previous page

	***Cc*Cry4a WT**								***Cc*Cry4a W369F**							
Peptide	**Deuteros**				**hdxstats**				**Deuteros**				**hdxstats**			
	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	DSM-DS	
490-495	5				x	x			1, 5, 30	1, 5			x	x		
495-505									15 s				x			
495-506									15 s, 1	5			x	x*		
496-505	1	1			x	x			15 s				x*			
496-506		1			x	x			15 s				x*			
506-522											5					
506-525				1												
506-529		30														
507-525		30														
507-529		30									5					
515-529		30	30													

Significant peptides showing LSM-LS differences for *Gg*Cry4a W369F were not found, neither using Deuteros nor hdxstats. The only DSM-DS difference found in Deuteros is the peptide 177-186. However, significant peptides found by hdxstats are 36-40, 36-41, 42-46, 47-57, four peptides covering the same region, and 161-176, 176-186, 177-186, three peptides covering the same protein region, as well as 220-228, 222-227, 222-228, another set of three peptides covering the same region. Furthermore, 315-322 came up as significant, three peptides spanning the region 404-426, 442-451, and seven peptides covering the region 495-517 towards the

C-terminal tail. It is surprising that so many peptides were found to be significant while there were none found in Deuteros. However, most of these peptides only had a FDR below 0.05 but not below 0.01. The two regions where peptides with a FDR below 0.01 were found, are between 222-228, just before the phosphate-binding loop, and between 495-505, towards the C-terminal end.

Summing up the findings for the LSM-LS and DSM-DS differences, there are four main regions of interest in which significant peptides were found repeatedly. 1. hdxstats found significant peptides in the N-terminal region that Deuteros did not find, consisting of several overlapping peptides for *Er*Cry4a W369F LSM-LS and for *Cc*Cry4a WT LSM-LS and DSM-DS, but not for the chicken cryptochromes. 2. Significant peptides in the region 161-176 were found in all three cryptochrome species, namely 167-175 for *Er*Cry4a W369F (hdxstats, LSM-LS), 161-173 for *Cc*Cry4a W369F (Deuteros, DSM-DS and hdxstats, DSM-DS), and 161-176 for *Gg*Cry4a W369F (hdxstats, DSM-DS) as well as more peptides covering the subsequent amino acids spanning a total region of 161-228. These peptides were not found in the WT and apart from *Er*Cry4a W369F mostly occurred for the DSM-DS difference. 3. Region 404-426 is also covered by significant peptides in all three bird cryptochrome species. For *Er*Cry4a W369F, it covers 404-412 (hdxstats, DSM-DS), for *Cc*Cry4a it is 413-425 (Deuteros, LSM-LS), and for *Gg*Cry4a W369F there are 404-412, 413-425, and 413-426 (all hdxstats, DSM-DS). The WTs of all three cryptochromes do not show any significant peptides in this region. 4. Finally, the C-terminal region repeatedly shows significant peptides. For *Er*Cry4a W369F, 505-526 is found to be significant in LSM-LS for both Deuteros and hdxstats, 506-527 and 516-527 are significant in LSM-LS within hdxstats. For *Cc*Cry4a WT, 506-525 is significant in for one time point (Deuteros, DSM-DS). For the W369F variant, 506-522 and 507-529 (Deuteros, LSM-LS) are significant for one time point each and for *Gg*Cry4a W369F several peptides span the region 495-517 (hdxstats, DSM-DS). Interestingly, we find again more significant peptides for the W369F mutant rather than the WT and most significant changes were found for

the LSM-LS difference in both migratory birds, but not in the chicken with one single exception for one time point for the quail WT protein.

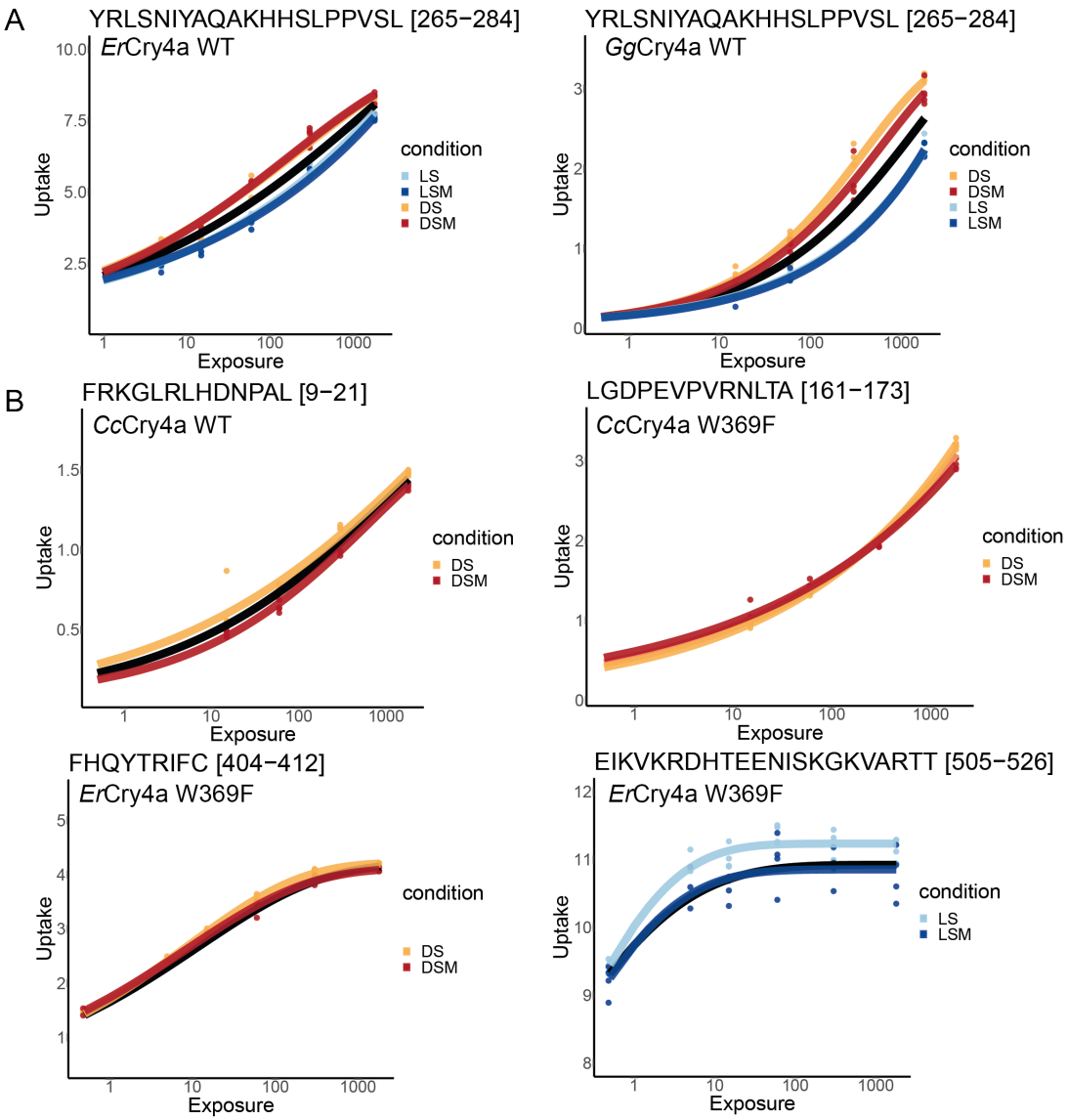

Figure 4.9: Example significant peptides found in hdxstats. (**A**) Example peptides with a significant difference for the LS-DS and LSM-DSM comparisons. (**B**) Four different significant peptides, each an example for a different protein regions where significant differences were found for the magnet-induced differences. Figure courtesy of Oliver Crook.

Some of the peptides discussed above are displayed in Figure 4.9 for illustration purposes. Panel A shows two peptides of two different proteins with a significant difference for the LS-DS and LSM-DSM comparisons. Panel B shows four different significant peptides, one from each of the protein regions where significant differences

were found for the magnet-induced differences.

With respect to the findings for the LS-DS and LSM-DSM differences, the hdxstats findings generally support the Deuteros results which are already discussed in detail above. An exception to the rule is the N-terminal region where there are no significant peptides in Deuteros, but several for all three bird cryptochrome species found by hdxstats. However, these are mostly not below a FDR value of 0.01. Generally, the trend seems to be that hdxstats usually finds more significant overlapping peptides than Deuteros. In a few instances, similar to the N-terminal region, there were significant peptides found in hdxstats where there were none or very few in Deuteros, most notably the regions 369-395 in *Er*Cry4a WT and 396-412 in *Gg*Cry4a WT.

Table 4.4: Significant peptides of *Gg*Cry4a WT and W369F. Significant peptides found by hdxstats are denoted as 'x'. Deuteros (p-value <0.01) and hdxstats results (empirical Bayes false discovery rate cutoff value <0.01; 0.01<value<0.05 marked with an asterisk) are compared. The Deuteros results list all time points in which the peptides were significant (listed in minutes unless otherwise indicated). If peptides were present in several charge states, the hdxstats results classify the result as <0.01 if this was true for at least one charge state. The table does not differentiate between charge states as some charge states were only present for some peptides (depending on the initial raw data). If there were no significant peptides found for any of the state differences, the respective columns are not displayed.

	***Gg*Cry4a WT**							***Gg*Cry4a W369F**					
Peptide	**Deuteros**			**hdxstats**				**Deuteros**			**hdxstats**		
	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	DSM-DS
9-15											x*		
9-21					x						x*		
9-22				x*	x*								
23-29						x*							

Continued on next page

Table 4.4 – continued from previous page

	***Gg*Cry4a WT**							***Gg*Cry4a W369F**					
Peptide	**Deuteros**			**hdxstats**				**Deuteros**			**hdxstats**		
	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	DSM-DS
36-40											x		x*
36-41											x*		x
36-46		5		x	x								
39-46				x	x						x	x	
41-46				x	x						x*	x*	
42-46											x*		x*
47-55											x*		
47-57				x*							x*		x*
61-67											x*		
75-81								30	30		x	x	
100-107					x*								
100-111								30					
100-112									30				
101-107												x	
113-123						x*							
124-128						x*	x*						

Continued on next page

Table 4.4 – continued from previous page

	***Gg*Cry4a WT**							***Gg*Cry4a W369F**					
Peptide	**Deuteros**			**hdxstats**				**Deuteros**			**hdxstats**		
	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	DSM-DS
134-142												x*	
143-150					x*								
143-154	5	5		x	x								
151-160				x*									
155-160				x	x						x	x	
161-174				x*									
161-176													x*
176-186													x*
177-186										30			x*
220-228													x*
222-227				x*									x
222-228											x*		x
229-243								1, 5	15 s, 1		x	x	
232-243	15 s, 5	1, 5		x	x			1, 5	15 s, 1, 5		x	x	
244-255								5, 30	5, 30		x	x	
244-256								30	5, 30		x	x	

Continued on next page

Table 4.4 – continued from previous page

	***Gg*Cry4a WT**							***Gg*Cry4a W369F**					
Peptide	**Deuteros**			**hdxstats**				**Deuteros**			**hdxstats**		
	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	DSM-DS
264-270	30	5, 30		x	x			5, 30	5, 30		x	x	
264-272	15 s, 5, 30	1, 30		x	x								
265-270	30	30		x	x	x		5, 30	5, 30		x	x	
265-272								1, 5, 30	15 s, 1, 5, 30		x	x	
265-284	15 s, 5, 30	1, 5, 30		x	x			1, 5, 30	15 s, 1, 5, 30		x	x	
271-284	5	1		x	x			1, 5	15 s, 1, 5, 30		x	x	
271-289	15 s, 5	1, 5, 30	30	x	x			1, 5, 30	15 s, 1, 5, 30		x	x	
272-289								1, 5, 30	15 s, 1, 5, 30		x	x	
273-284								1, 5	15 s, 1, 5		x	x	
273-289	15 s, 1, 5, 30	1		x	x			1, 5, 30	15 s, 1, 5, 30		x	x	

Continued on next page

Table 4.4 – continued from previous page

	***Gg*Cry4a WT**							***Gg*Cry4a W369F**					
Peptide	**Deuteros**			**hdxstats**				**Deuteros**			**hdxstats**		
	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	DSM-DS
285-289								30	30		x	x	
289-294					x*								
295-304				x*							x	x	
299-304												x*	
305-314								30	30		x	x	
315-322													x*
343-361	30			x	x			30			x		
343-362				x	x	x*					x		
346-361								5, 30			x		
346-362				x	x	x*					x		
362-368								30	30		x	x	
362-369								30	30		x	x	
363-368						x	x						
369-375						x*							
383-387				x	x	x*							
384-388								1	15 s, 1		x	x	

Continued on next page

Table 4.4 – continued from previous page

	***Gg*Cry4a WT**							***Gg*Cry4a W369F**					
Peptide	**Deuteros**			**hdxstats**				**Deuteros**			**hdxstats**		
	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	DSM-DS
388-394											x		
388-395				x	x	x	x*				x*		
389-395											x*		
396-400					x*			30	30		x	x	
401-412		1, 5		x	x								
404-412				x	x			1	15 s, 1		x	x	x*
406-412									15 s		x	x	
413-425								1			x	x	x*
413-426								1, 5	5, 30		x	x	x*
417-425								30					
417-426								30			x	x	
426-434											x*		
427-434											x*		
442-451											x*		x*
458-481	30			x									
464-481								1, 5, 30	1, 5, 30		x	x	

Continued on next page

Table 4.4 – continued from previous page

	***Gg*Cry4a WT**							***Gg*Cry4a W369F**					
Peptide	**Deuteros**			**hdxstats**				**Deuteros**			**hdxstats**		
	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	LSM-LS	DSM-DS	LS-DS	LSM-DSM	DSM-DS	LS-DS	LSM-DSM	DSM-DS
471-481								1, 5, 30	1, 5, 30		x	x	
482-495	5, 30	1, 30		x	x			1, 5, 30	15 s, 1, 5, 30		x	x	
495-505									15 s		x	x	x
496-505	15 s	1		x	x				15 s		x	x	x
496-506	15 s	1		x	x				15 s				x*
499-506									15 s		x	x	x*
506-516													x*
507-516													x*
507-517													x*

4.2.4 Conclusions

This chapter presented a large scale comparison of seven different cryptochrome 4a proteins. The aim was to investigate protein areas that might undergo conformational changes by measuring changes in surface accessibility when comparing different initial conditions with each other involving the presence and absence of blue light and permanent magnets. For this, the cryptochromes of two migratory and of one non-migratory avian species, European robin, quail, and chicken, respectively, were compared and the results underwent thorough statistical analysis using several different approaches.

Ten distinct regions where changes in surface accessibility could be observed were identified. All of these areas are situated in one half of the protein which corresponds to the lower half of the protein displayed in Figure 4.1. Several of these regions stand out as they are covered by multiple overlapping, significant peptides in almost all state comparisons for all proteins. These are the PBL, the PM, the region between residues 404-426, α22 and the residues after this alpha-helix towards the C-terminal end of this protein. These were also the areas found by Kattnig *et al.* in 2018 to undergo reorganisation when performing molecular dynamics simulations on light-excited *Er*Cry4a WT [8]. Interestingly, residues 404-426 include residue 412 which was found to be important for dimerisation of *Er*Cry4a [17] (see also chapter 5). Since the experiments were conducted on monomeric proteins, it is interesting to note that mostly protection was observed, but no de-protection in adjacent areas which one might expect to find. One reason, as mentioned earlier, could be that the PBL and the PM as well as region 404-426 and part of the C-terminal tail cover each other up. Given that they largely consist of loops and and highly solvent exposed, probably rather flexible areas, this would not require large areas to show any de-protection. There is a slight chance that the proteins might form non-covalent dimers (see also chapter 5). However, no significant increase in dimerisation seemed to occur during up to 30 min of native MS measurement (data not displayed) which makes it unlikely that the very consistent and large light-induced differences might be a result of dimerisation.

GgCry4a W369F stands out as the ΔRFU changes observed for this protein for the light-induced changes have very high values, especially in the region of the PM. This protein as well as the other two W369F mutants also display comparatively larger changes in α22 than their WT counterparts as well as protection events in their mutation sites (except for *Cc*Cry4a), reassuringly showing that this method is sensitive to small changes in the protein structure. Further differences between the W369F mutant and the corresponding WT have been found which taken together show that there is generally a greater variability within the W369F mutant with

respect to their numbers of overlapping peptides, ΔRFU values, and consistency of significant protein areas across labelling time points.

Another reason these experiments were conducted, was to find magnet-induced differences. Several peptides were found to be significant in the DSM-DS comparison. However, since without any radical formation no additional effects of a magnet on the sample were expected, these peptides are considered to be outliers. Given that the LSM-LS comparison also only yielded a few significant peptides, the significance of these findings need to be considered cautiously as these could be constituted at least in parts of outliers as well. Significant peptides found in Deuteros show protection for *Er*Cry4a W369F and *Cc*Cry4a W369F in and after α23 (505-529) for the LSM-LS comparison which is not the case for the chicken cryptochrome. This could cautiously be interpreted as a difference between migratory and non-migratory bird cryptochromes, but given that there are not many overlapping peptides strengthening this case and remembering that outliers might be among these peptides, this is a rather weak conclusion. Additionally, this finding was only backed up by the results found in hdxstats for *Er*Cry4a W369F, but not for *Cc*Cry4a W369F where hdxstats found no significant peptides for the LSM-LS difference whatsoever. In general, after thorough analysis of the data, there are some weak trends, but no conclusions that can be made with confidence with respect to the magnet-induced changes in surface accessibility. The most notable differences are probably several overlapping significant peptides in the C-terminal region of *Er*Cry4a W369F for the LSM-LS comparison and for *Gg*Cry4a W369F for the DSM-DS comparison. However, it is interesting to note that the magnet-induced significant peptides can roughly be divided into four different regions, namely the N-terminus, amino acid regions 161-176 (just after α6) and 404-426, and the C-terminal tail of which the latter two also regularly come up as light-induced changes.

Concluding, our detailed analysis revealed multiple differences between the different bird cryptochromes investigated, but also several notable similarities which are summarised in the following:

- the FAD cofactor is necessary to observe light-induced changes in uptake

- all bird cryptochromes investigated display multiple regions which become less solvent-accessible upon exposure to blue light
- the most notable regions that become more protected in blue light are the PBL, the PM, the region covering residues 404-426, and the CTT - one possible explanation for this result being that the PBL and the PM as well as region 404-426 and the CTT cover each other after undergoing conformational changes
- only single peptides, not whole regions, were shown to become more solvent-exposed
- the W369F mutants display comparatively larger changes in α22 than their WT counterparts; they also display significant changes in the region of mutation at long labelling time points
- the *Gg*Cry4a W369F mutant displays exceptionally large ΔRFU changes, especially in the PM
- uptake differences observed for the CTT, especially for α23 and beyond, need to be interpreted carefully given that the region is already highly deuterated starting at the earliest labelling time points, making the findings of significant changes in uptake less reliable
- the magnet-induced differences offer weak trends, but no conclusive findings, and therefore cannot distinguish between migratory and non-migratory cryptochrome species

4.3 Outlook and Interaction partner studies

Further experiments based on the findings in these data sets are currently being undertaken by Shah. An eighth HDX-MS experiment has already been performed by her on *Er*Cry4a C317R+W369F+Y319F and is currently being analysed. This mutant was chosen as in spectroscopic measurements the *Er*Cry4a W369F+Y319F

double mutant had proven to be more stable than the WT. The C317R mutation was introduced as this removes an amino acid suspected to cause instability in comparison to the more stable *Gg*Cry4a WT which has an arginine in the same position. The hope is to deepen our knowledge of changes in solvent accessibility changes of cryptochromes by investigating a protein known to be very stable. Furthermore, Shah is working on pulsed labelling HDX-MS experiments where *Er*Cry4a WT and *Cc*Cry4a WT are exposed to blue light and subsequently to darkness in order to probe the proteins for any backwards reaction.

As mentioned in the introduction to this chapter, the aim was originally to follow up on the findings presented here by adding interaction partner studies using the same HDX-MS approach in order to elucidate which parts of the cryptochromes interact with which part of the interaction partners. This could have led to new insights into how the observed conformational changes might impact or even be necessary for the downstream communication pathway connecting the perception of the Earth's magnetic field inclination to navigational behaviour. A collaboration for this purpose was established with Karl Koch's research group at the Carl von Ossietzky University of Oldenburg in Germany who were able to express one of the putative interaction partners [2] found in 2020 by Wu *et al.* [1], GNAT2. Additionally, we also received the long-wavelength opsin loop 2 (LWO2) as a part of another putative interaction partner of Cry4a and a scrambled version of this loop as a control. Using native MS, we were able to obtain initial, promising results showing interactions between Cry4a and the loop (data not shown here). So far, the interaction between GNAT2 and Cry4a has been proven to be more difficult to observe, especially given a short activity time frame of GNAT2 that lasts only about a week. Up until now, efforts to improve the conditions for the observation of an interaction included varying the salt concentrations and the pH of the spraying buffer, as well as using different incubation times ranging from a few seconds to 15 min, applying different light conditions (low light (complete darkness was impossible due to the layout of the current laboratory setup) and blue light), and measuring after incubation at room temperature and on ice. However, none of the different

conditions probe so far had a significant impact on the outcome of the experiment. Time was too short to push this project beyond initial native MS experiments and HDX-MS test runs and the work will subsequently be continued by Shah.

4.4 Materials and Methods

4.4.1 Materials

Reagents and chemicals were purchased from Sigma Aldrich (Gillingham, UK). Proteins were expressed in and purified from *E. coli* by our collaborators in the group of Henrik Mouritsen at the Carl von Ossietzky University of Oldenburg, Germany. Unless indicated otherwise, all samples were shipped on dry ice in 20 mM Tris, 250 mM NaCl, 10 mM BME, and 20% glycerol and kept at -80°C until further use.

The equilibration buffer used consisted of 5 mM potassium phosphate (H_2O), pH 7.0, the labelling buffer consisted of 5 mM potassium phosphate (D_2O), pD 7.0, and the quench buffer consisted of 50 mM potassium phosphate (H_2O), pH 1.9.

4.4.2 HDX-MS

The proteins were diluted down to 10-16 µL and essential protein handling was performed under dim red light conditions. All experiments were performed in a dark room. All proteins were equilibrated for at least 20 min in a temperature controlled chamber at 20°C under different environmental conditions prior to labelling: the light state of the protein was induced by continuous illumination with 450 nm blue LEDs at powers between 0.4-0.7 mW cm^{-2}, the dark state of the protein was induced by keeping it in complete darkness and the magnet states were induced by inserting neodymium magnets of 160 mT strength below the samples in a custom-built vial holder. The magnet strength directly applied to the proteins lies between ∼18-25 mT with a distance of approximately ∼2-3 mm (see Figure 3.9 in chapter 3 for details).

Proteins were labelled by adding 55 µL of the labelling buffer at room temperature to 5 µL of sample resulting in a labelling solution of ∼92% D_2O. After varying labelling times the reactions were quenched by adding an equal volume of a quench

solution that had been kept at 0°C followed by immediate injection of the quenched solution into the HDX manager (Waters, Milford, MA, USA). Subsequently the solution was kept at 0°C until reaching the pepsin column Enzymate Pepsin (Waters) for digestion where it was held at 15°C using an isocratic H_2O (0.1 % v/v) formic acid solution at a flow rate of 100 µL/min. This was followed by 2 min of product collection and de-salting on a BEH C_{18} trapping column (Waters) and separation on a reversed-phase analytical column (Waters) at 0°C using a flow rate of 40 µL/min.

The elution gradient for separation started with 97 % H_2O and 3 % MeCN changing to 65 % H_2O and 35 % MeCN within 7 minutes followed by a gradient going from 65% to 5% H_2O within minutes 7-8 and held at 5% H_2O for minutes 8-10. All peptides were electrosprayed into a Synapt G2Si Q-ToF instrument (Waters) for mass analysis and all sample handling was performed manually. The instrument parameters were as follows: capillary at 2.8 kV, sample cone at 100 V, source offset at 30 V, source temperature at 80°C, and the desolvation temperature at 150°C. LeuEnk was used for internal calibration and its m/z acquired every 30 s.

A cleaning cycle with a saw-tooth gradient was performed in between each labelling protein run using a washing solution of 4 % acetonitrile (ACN) and 0.1 % formic acid in water with 0.5 M guanidinium hydrochloride (GdnHCl). Intermittent blank runs were injected to ensure the carryover between different protein runs to be minimal.

The three small red lights used to illuminate the work bench and the blue LEDs used to illuminate the samples when necessary were tested for their emission spectra to avoid unwanted overlap of wavelengths. The recorded emission spectra displayed in Figure 3.10 of chapter 3.

4.4.3 Native MS

The same samples as used for the HDX-MS experiments were used for the native MS measurements taken from a separately sent vial to avoid thawing the samples unnecessarily. The samples were kept at -80°C until they were thawed on ice and buffer-exchanged into 200 mM ammonium acetate, pH 8, for the native MS

measurements using Zeba^{TM} Micro Spin desalting columns with a molecular weight exclusion limit of 40 kDa (ThermoFisher Scientific).

Measurements were performed in Oxford using in-house gold-plated capillaries on a Q Exactiv^{TM} mass spectrometer in positive ion mode. The samples were sprayed at approximately 8 µM using nanoflow electrospray with a source temperature of 150°C and a capillary voltage of 0.9-1.2 kV. The higher-energy C-trap dissociation (HCD) cell voltage was 5 V, in-source trapping was set to -200 V to help with the dissociation of small ion adducts, and the noise threshold parameter was 3. Ion transfer optics and voltage gradients throughout the instruments were adjusted for optimum protein transmission.

4.4.4 Data analysis

Peptides were identified by data-independent MS/MS (MS^E) analysis and subsequent database searching in Protein Lynx Global server 3.0 (PLGS; Waters). Peptide identifications were filtered according to fragmentation quality (minimum fragmentation products per amino acid: 0.2), mass accuracy (maximum [MH]+ error: 5 ppm), maximum sequence length (25 amino acids), and reproducibility (peptides identified in at least three MS^E repeats) before their integration into HDX analysis. HDX-MS data were processed in DynamX 3.0 (Waters), and all automated peptide assignments were manually verified, with noisy and overlapping spectra discarded. Deuteros 2.0 [5, 13], HD-eXplosion [15], and hdxstats [16] were used for the statistical data analysis of the HDX-MS data. The PyMOL 2.4.3 software[18] (Schrödinger, NY, USA) as well as the UCSF Chimera software[19] (version 1.14rc) were used for 3D data visualisation.

References

1. Wu, H., Scholten, A., Einwich, A., Mouritsen, H. & Koch, K.-W. Protein-protein interaction of the putative magnetoreceptor cryptochrome 4 expressed in the avian retina. *Scientific Reports* **10,** 7364. http://www.nature.com/articles/s41598-020-64429-y%20https://www.nature.com/articles/s41598-020-64429-y (2020).

2. Görtemaker, K. *et al.* Direct Interaction of Avian Cryptochrome 4 with a Cone Specific G-Protein. *Cells* **11,** 2043. https://www.mdpi.com/2073-4409/11/13/2043 (2022).

3. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nature Methods* **19,** 679–682. https://www.nature.com/articles/s41592-022-01488-1 (2022).

4. Madeira, F. *et al.* Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research* **50,** W276–W279. https://academic.oup.com/nar/article/50/W1/W276/6567472 (2022).

5. Lau, A. M., Claesen, J., Hansen, K. & Politis, A. Deuteros 2.0: peptide-level significance testing of data from hydrogen deuterium exchange mass spectrometry. *Bioinformatics* **37** (ed Elofsson, A.) 270–272. https://academic.oup.com/bioinformatics/article/37/2/270/5877424 (2021).

6. Wales, T. E. & Engen, J. R. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrometry Reviews* **25,** 158–170. https://onlinelibrary.wiley.com/doi/10.1002/mas.20064 (2006).

7. Houde, D., Berkowitz, S. A. & Engen, J. R. The Utility of Hydrogen/Deuterium Exchange Mass Spectrometry in Biopharmaceutical Comparability Studies. *Journal of Pharmaceutical Sciences* **100,** 2071–2086. https://linkinghub.elsevier.com/retrieve/pii/S0022354915320955 (2011).

8. Kattnig, D. R., Nielsen, C. & Solov'yov, I. A. Molecular dynamics simulations disclose early stages of the photo-activation of cryptochrome 4. *New Journal of Physics* **20,** 083018. https://iopscience.iop.org/article/10.1088/1367-2630/aad70f (2018).

9. Inglis, A. J. *et al.* Activation of GCN2 by the ribosomal P-stalk. *Proceedings of the National Academy of Sciences* **116,** 4946–4954. https://pnas.org/doi/full/10.1073/pnas.1813352116 (2019).

10. Ravenhill, B. J. *et al.* The Cargo Receptor NDP52 Initiates Selective Autophagy by Recruiting the ULK Complex to Cytosol-Invading Bacteria. *Molecular Cell* **74,** 320–329.e6. https://doi.org/10.1016/j.molcel.2019.01.041%20https://linkinghub.elsevier.com/retrieve/pii/S1097276519300619 (2019).

11. Ohashi, Y. *et al.* Membrane characteristics tune activities of endosomal and autophagic human VPS34 complexes. *eLife* **9,** 1–28. https://elifesciences.org/articles/58281 (2020).

12. Xu, J. *et al.* Magnetic sensitivity of cryptochrome 4 from a migratory songbird. *Nature* **594,** 535–540. http://dx.doi.org/10.1038/s41586-021-03618-9 (2021).

13. Lau, A. M. C., Ahdash, Z., Martens, C. & Politis, A. Deuteros: software for rapid analysis and visualization of data from differential hydrogen deuterium exchange-mass spectrometry. *Bioinformatics* **35** (ed Valencia, A.) 3171–3173. https://academic.oup.com/bioinformatics/article/35/17/3171/5288775 (2019).

14. Schuhmann, F., Kattnig, D. R. & Solov'yov, I. A. Exploring Post-activation Conformational Changes in Pigeon Cryptochrome 4. *Journal of Physical Chemistry B* **125,** 9652–9659 (2021).

15. Zhang, N., Yu, X., Zhang, X. & D'Arcy, S. HD-eXplosion: visualization of hydrogen–deuterium exchange data as chiclet and volcano plots with statistical filtering. *Bioinformatics* **37** (ed Yann, P.) 1926–1927. https://academic.oup.com/bioinformatics/article/37/13/1926/5929686 (2021).

16. Crook, O. M., Chung, C.-w. & Deane, C. M. Empirical Bayes functional models for hydrogen deuterium exchange mass spectrometry. *Communications Biology* **5,** 588. https://www.nature.com/articles/s42003-022-03517-3 (2022).

17. Hanić, M. *et al.* Dimerization of European Robin Cryptochrome 4a. *The Journal of Physical Chemistry B* **127,** 6251–6264. https://pubs.acs.org/doi/10.1021/acs.jpcb.3c01305 (2023).

18. Schrödinger, L. & DeLano, W. *The PyMOL Molecular Graphics System* 2022. http://www.pymol.org/pymol.

19. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25,** 1605–1612. https://onlinelibrary.wiley.com/doi/10.1002/jcc.20084 (2004).

You're a difficult equation with a knack for heart evasion
Will you listen to my proof or will you add another page on?
It appears to me the graph has come and stolen all the laughs
It appears to me the pen is over-analyzed again

— from Sia's *Academia*

5

Dimerisation of *Er*Cry4a

Contents

5.1 Introduction 176
5.2 Results and Discussion 176
5.2.1 Covalent dimerisation of *Er*Cry4a 176
5.2.2 Chemical cross-linking followed by mass spectrometric analysis 183
5.2.3 Focus on C317 and C412 184
5.2.4 Dimerisation under different light conditions 186
5.2.5 The search for the interaction interface 192
5.3 Conclusion 201
5.4 Materials and Methods 203
5.4.1 Materials 203
5.4.2 Native MS 203
5.4.3 Mass photometry 204
5.4.4 SDS-PAGE 204
5.4.5 Chemical cross-linking and mass spectrometric analysis 205
5.4.6 DEER 206
References 207

5.1 Introduction

Dimerisation and higher order oligomerisation is a common theme in the world of protein structure-function analysis. In this chapter the homo-dimerisation of robin cryptochrome 4 is explored and its relevance with respect to its role as the putative magnetoreceptor is discussed using results from a combined experimental and computational investigation. The experimental studies using native mass spectrometry, SDS-PAGE, mass spectrometric analysis of disulphide bonds, chemical cross-linking and photometric measurements (the latter performed by Katharina Görtemaker) were complemented by the theoretical studies of possible covalently and non-covalently bound dimers using computational modelling and molecular dynamics simulations, performed by Maja Hanić and Lewis M. Antill. The results presented in this chapter are the product of a collaborative effort which has recently been published [1].

5.2 Results and Discussion

5.2.1 Covalent dimerisation of *Er*Cry4a

Initially, native MS, mass photometry and sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) were used to investigate the dimeric nature of *Er*Cry4a as displayed in Fig. 5.1. Native MS consistently showed the presence of dimeric structures at micromolar protein concentrations, as shown by the example given in Fig. 5.1A where the major peaks correspond to the monomer, and minor peaks in the range m/z = 5000-7000 to the dimer. Even at the nanomolar protein concentrations used for mass photometry measurements dimers could still be observed as shown in 5.1B displaying a major component in the range of 60-80 kDa (86%) attributed to monomers and a minor component between 120-160 kDa (11%) attributed to dimers. Lowering the concentration did not seem to change the monomer-dimer ratio significantly which suggests that the two monomer units are strongly bound via a specific interaction and not an artefact resulting from nanoelectrospray ionisation. Different amounts of reducing agents reduced the

number of dimers and higher order oligomers seen in a denaturing SDS-PAGE gel in 5.1C, suggesting the presence of covalently linked monomers via cysteine bonds. The complete gel is shown in Fig. 5.2A.

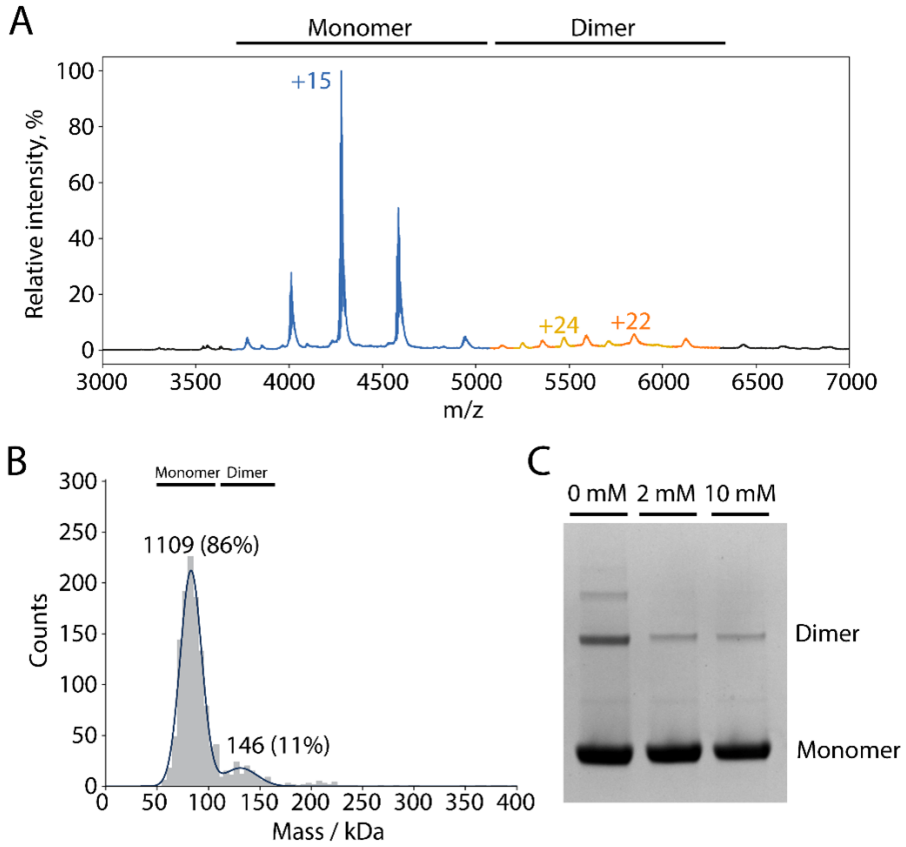

Figure 5.1: Dimerisation in *Er*Cry4a. (**A**) Native mass spectrum of *Er*Cry4a. The peaks of the monomer region are indicated in blue (molecular mass 64.197 ± 0.029 kDa) and the peaks of the dimer region in orange and yellow (molecular masses ∼123 kDa and ∼131 kDa). The expected mass of an *Er*Cry4a WT monomer including FAD and a His-tag is 64.034 kDa. The differences between the detected and the expected masses are due to salt ions and buffer components. The finding of a second dimer mass, lower than the expected mass, could be due to minor truncations and is only observed in some samples. The observation here of stable dimers is consistent with Fig. S13 of Xu et al [2]. The sample used for this measurement contained 10 mM BME during shipment to prevent higher order oligomerisation. The result depicted reflects a typical mass spectrum of a cryptochrome. (**B**) Mass photometry of 44 nM *Er*Cry4a. The same sample as in (A) but shipped with 2 mM DTT was used. The numbers above the peaks are the numbers of counts. (**C**) SDS-PAGE gel of *Er*Cry4a samples without a His-tag (the same as used for XL-MS) that had been incubated with various amounts of DTT.

The same figure also shows the results of an additional investigation of the strength of the interaction by using native tandem MS in order to isolate the

monomer and dimer peaks separately of two proteins available at the time, namely *Er*Cry4a WT and *Er*Cry4a C317S, to compare their stability upon exposure to high energy collisional dissociation (HCD) conditions. The dimer does not dissociate, even at very high HCD values, whereas the monomer can be seen to slowly fall apart at lower HCD values (Fig. 5.2D & E). Fig. 5.2E displays the isolation of a pentamer peak of the non-covalently bound C-reactive protein (CRP) which was used as a control and dissociates into its subunits at much lower HCD values than those applied to both the monomer and dimer of the cryptochromes. This points towards a comparatively strong interaction between two cryptochrome subunits within a dimer.

Molecular dynamics simulations of a previously used *Er*Cry4a construct [2] including its cofactor FAD were used by Maja Hanić for the analysis of the solvent accessible surface area (SASA) [4] using the online tool GETAREA [5]. The SASA values per residue were subsequently divided by the SASA value of the amino acid in question (X) in a Gly-X-Gly tripeptide (which had previously been constructed in Pep McConst [6]), using a similar approach as published elsewhere [7]. The result in percent is displayed in Fig. 5.3 and revealed that six out of the eleven present cysteines in the protein are solvent-exposed, i.e. have a sidechain accessibility greater than 20% with the other five being below 7%. These residues are C68, C73, C116, C189, C317, and C412 and therefore assumed to possibly be involved in the formation of disulphide bonds.

Interesting in this context was the analysis of the cysteine solvent accessibility of different cryptochromes in solution performed by Katharina Görtemaker. The cysteine accessibility was assessed by measuring the absorption of the product formed by treating the protein solution with the Ellmann's reagent 5,5-dithio-bis-(2-nitrobenzoic acid) (DTNB). More information on the experimental procedure can be found in Hanić *et al.*[1].

The results displayed in Table 5.1 show that approximately five cysteines are accessible for *Er*Cry4a WT, four for each of the C317A and C412A mutants and three for the C116A+C317A mutant. These results suggest that the residues C116, C317 and C412 are indeed significantly solvent-exposed. Although this

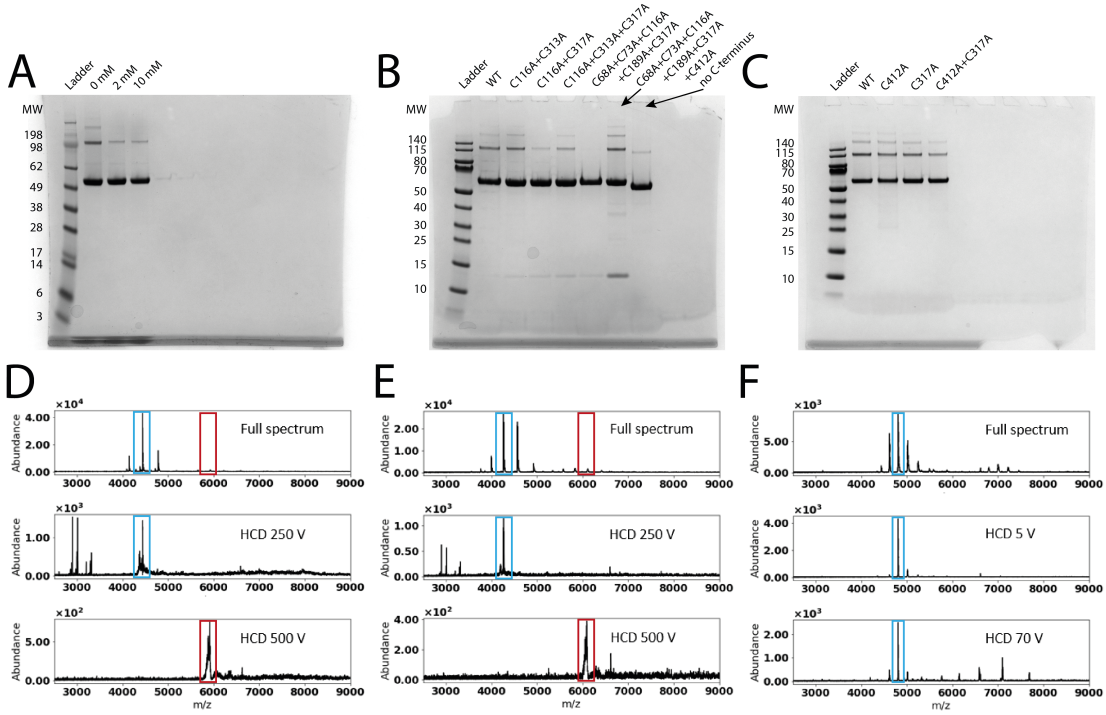

Figure 5.2: Close up on dimerisation. (**A**)-(**C**) Full denaturing SDS-PAGE gels run for the different *Er*Cry4a proteins. Proteins are as indicated at the tops of the columns. All leftmost columns display protein ladders. In (**A**) the SeeBlueT^{TM} Plus2 Pre-stained Protein Standard ladder (Invitrogen) was used and in (**B**) and (**C**) the PageRuler Prestained Protein Ladder (Thermo Scientific) was used with the molecular weight (MW) mass markers in kDa as indicated. (**A**) shows the full SDS-PAGE gel partially displayed in Fig. 5.1C and (**C**) the one partially displayed in Fig. 5.5C. (**D**)-(**F**) Native MS in combination with tandem MS was used to isolate a monomer and a dimer peak of *Er*Cry4a WT (**D**) (without its His-tag) and *Er*Cry4a C317S (**E**) (with His-tag; shipped in 10 mM BME to prevent higher order oligomerisation during transport) to compare their stability upon exposure to high HCD (high energy collisional dissociation) energies. The first rows of both (**D**) and (**E**) show the full spectra, the second the isolated dimer peaks and the third the isolated monomer peaks. HCD values were applied as indicated in the spectra. The spectra displayed for comparison in (**F**) are of C-reactive protein (CRP), a pentameric protein known to be non-covalently bound [3]. The first row shows the full spectrum, the second an isolated pentamer peak and the third shows the isolated pentamer peak after applying an HCD value of 70 V, displaying how it dissociates into smaller subunits at comparatively low HCD energies. The *Er*Cry4a dimers still did not dissociate when the highest HCD energies possible on the instrumental setup (500 V) were applied. The coloured rectangles show which peaks were isolated using tandem MS.

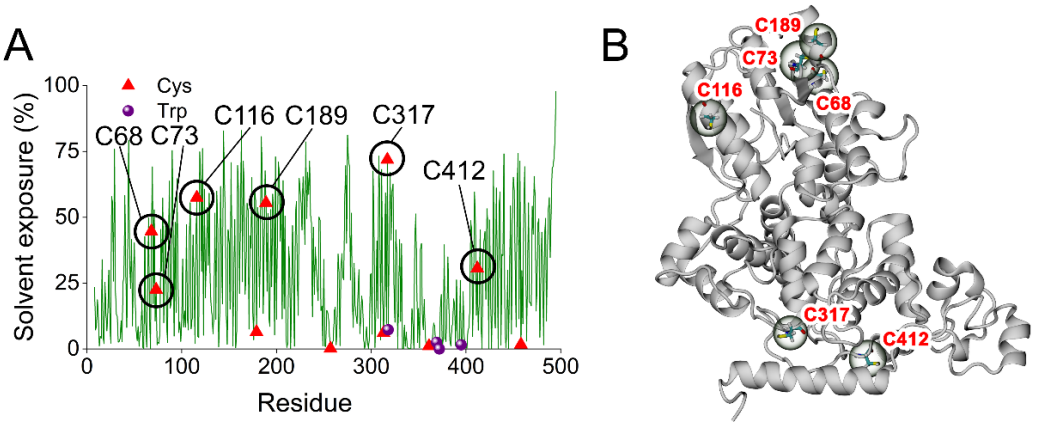

Figure 5.3: Solvent accessible cysteine residues. (**A**) Solvent exposures of residues in monomeric WT *Er*Cry4a with cysteines (red triangles) indicated. Also shown (purple spheres) are the components of the Trp-tetrad: (W318, W369, W372, W395). (**B**) *Er*Cry4a secondary structural elements with the six most exposed cysteines indicated. Figure courtesy of Maja Hanić.

means that they are in principle available for disulphide bonding, it should be added that due to the small size of DTNB, it would be possible that the reaction with the Ellman's reagent might occur despite the residue not being sufficiently solvent-exposed for disulphide bonding.

Table 5.1: TNB absorbance measurements (at 412 nm) and average number of accessible cysteines in WT and mutant forms of *Er*Cry4a. Table courtesy of Katharina Görtemaker.

Protein	A_{412}	Accessible Cysteines
WT	0.30 ± 0.01	4.78 ± 0.13
C317A	0.27 ± 0.01	4.23 ± 0.10
C412A	0.27 ± 0.00	4.16 ± 0.05
C116A+C317A	0.17 ± 0.01	3.30 ± 0.11

Alongside the acquisiton of the experimental results a theoretical analysis of different possible covalently bound dimers using the molecular docking tools ZDOCK [8–10] and M-ZDOCK [11] and molecular dynamics (MD) simulations using NAMD [12, 13] were performed by Maja Hanić. For some of the docking approaches, no contact residues were defined, for others the most solvent-exposed residues were used as predefined docking sites in ZDOCK. All docked structures with their

cysteine residues in close enough contact for them to build disulphide bridges were chosen for subsequent MD simulations.

As subsequent experiments displayed in Fig. 5.4 in section 5.2.2 below revealed, residue C412 seems to be of importance for the covalent dimerisation process and was therefore further investigated by imposing an artificial constraint on the C412-C412 bond to enable these residues to move closer together to create additional configurations that were subsequently simulated.

Molecular modelling of non-covalently bound dimers was performed by Lewis M. Antill. For this, a full length *Er*Cry4a homology model based on pigeon cryptochrome 4 (PDB code: 6PU0) [14] was created with the missing C-terminal tail (CTT) consisting of residues 498-527 being created with I-TASSER [15–18]. ZDOCK was again used as the docking tool. Inspired by a non-covalent, crystallographic dimer of *Mm*Cry2 (PDB: 6KX8) [19] which is structurally the closest protein to *Er*Cry4a for which a dimer crystal structure has been reported and also considering that dimer interfaces tend to have a high degree of conservation in evolutionary related proteins [20], another non-covalent dimer was constructed by Maja Hanić in which its subunits were structurally aligned with the *Mm*Cry2 crystal structure.

Maja Hanić's and Lewis M. Antill's work led to the results displayed in Table 5.2 showing the lowest non-bonded monomer-monomer interaction energies (more details on their methods and further results can be found in Hanić *et al.* [1]). These are two structures with C412-C412 disulphides (cov^{412A} and cov^{317B}), two with C317-C317 bonds (cov^{317A} and cov^{317B}), two non-covalent structures (ncov^{A} and ncov^{M}), and one with two disulphide bridges (cov^{D}; C116-C313 and C313-C116). ncov^{M} is the structure based on the crystallographic dimer of *Mm*Cry2.

ncov^{A} and cov^{412A} are the most stable structures found as they have the lowest E_{tot} values. The lower the radius of gyration, R_g, which is a measure of the compactness of the structure, the stronger the binding is expected to be, which is reflected in the two non-covalent structures listed in the table. These are also the two structures with the most hydrogen bonds with ncov^{A} also having the highest number of salt bridges. The root-mean square deviation (RMSD) is another measure

Table 5.2: The seven simulated and analysed *Er*Cry4a dimers with the lowest non-bonded monomer-monomer interaction energies, E_{tot}. The data displayed were acquired by and used with the permission of Maja Hanić and Lewis M. Antill. (n)cov^{N}n - (non-)covalent structures where N is the sequence number of the cysteine residue that forms the disulphide bond and n the label for a particular dimer (the full list of dimers can be found in Hanić *et al.* and the corresponding Supplementary Material [1]), E_{tot} - internal potential energy (includes non-covalent van der Waals and electrostatic interactions but does not include the ∼60 kcal mol^{-1} expected for a disulphide bond [21]), R_{g} - radius of gyration reflecting the compactness of the structure, A_{IS} - interaction surface area (difference between the solvent accessible surface areas of the separate monomers and that of the dimer).

Dimer	$-E_{\mathrm{tot}}/kcalmol^{-1}$	R_{g} / Å	RMSD / Å	RMSF / Å	A_{IS}/Å^2	Hydrogen bonds	Salt bridges
$ncov^{A}$	857±85	35.2±0.3	3.5±0.4	2.2±1.0	2762±292	106	29
cov^{412A}	526±83	41.2±0.2	2.5±0.4	5.7±1.3	1840±115	31	13
$ncov^{M}$	505±163	33.0±0.3	3.4±1.4	2.3±1.0	2442±284	111	12
cov^{412B}	437±127	41.4±0.7	6.2±2.8	5.4±1.5	1040±310	49	12
cov^{317A}	186±103	38.3±0.3	4.0±0.6	1.8±0.9	1082±220	64	10
cov^{D}	178±44	33.0±0.2	2.5±0.3	1.3±0.7	2140±330	47	3
cov^{317B}	171±73	38.5±0.2	3.8±0.8	15.6±4.1	1070±110	59	12

of protein stability as it indicates the deviation of the protein backbone from a reference structure, usually the first frame of a MD production run [22]. Apart from cov^{412B}, all listed structures have relatively low RMSD values below 4 Å. The root-mean square fluctuation (RMSF), a value to assess the degree of internal mobility by comparing the positions of each carbon atom at every simulation step to its position in the average structure, reveals comparable mobility for all structures listed except for cov^{317B} which is less rigid. Larger interaction surface area values A_{IS} seem to correlate roughly with E_{tot}. A display of all seven structures mentioned in Table 5.2 can be found in the Appendix (see Figures A.36-A.39). Further details and full lists of all structures obtained during docking procedures can be found in Hanić *et al.* and the corresponding Supplementary Material [1].

As part of the theoretical investigations into possible dimeric structures, the selection criterion that sulphur atoms should be no more than 6 Å apart during the docking simulations was applied. During this process three residues were identified as potentially being involved in disulphide bridges, namely residues C116, C313, and C317, out of which the first and the last have the two highest calculated surface exposures (more details in S18 of Hanić *et al.* [1]). Following this finding, the

collaborating group of Henrik Mouritsen at the Carl von Ossietzky University of Oldenburg expressed the following mutants: C317A, C116A+C313A, C116A+C317A, C116A+C313A+C317A, and C68A+C73A+C116A+C189A+C317A (with five of the six most surface accessible cysteines replaced by alanines). However, Fig. 5.2B shows that all four mutants still oligomerise. Therefore, if these residues are involved in dimerisation, they are not the only ones.

5.2.2 Chemical cross-linking followed by mass spectrometric analysis

Chemical cross-linking followed by mass spectrometric analysis as well as mass spectrometric analysis of the non-cross-linked dimer were performed to investigate the interface between monomer units. For the former, the bifunctional compounds DSSO [23] and DSBU [24] were used to cross-link the primary amino groups of lysine residues in close proximity, via bis(N-hydroxysuccinimide) ester groups. The linkers contain a MS cleavable moiety, therefore subsequent tandem MS analysis using gas-fragmentation in the HCD cell was able to identify the cross-linked peptides and thereby the involved lysines with α-carbons separated by up to $\sim$27 Å. More specifically, based on linker lengths of 12.5 Å and 10.3 Å maximum distances between two α-carbon atoms of 26.9 Å and 23.5 Å can be measured using DSBU and DSSO, respectively. The results are displayed in Table 5.3 and visualised in Fig. 5.4. Monomer and dimer units were investigated separately and their results therefore assumed to be intramolecular and intermolecular cross-links, respectively. Three out of the four disulphide bonds identified (C361-C458, C412-C361, C412-C412, and C412-C458) involved C412 with only the C412-C412 being unequivocally intermolecular. This is due to the difficulty to distinguish between the options of intra- and intermolecular dimers in a homodimer. The analysis of the artificially cross-linked proteins yielded five lysine residues involved in the formation of cross-links. The cross-link involving two K152 residues is consistent with a C412-C412 disulphide bridge because K152 and C412 are situated on the same side of the protein as displayed in Fig. 5.4B. The links from K234 to K429 and from K152 to

K507 were found in both the monomer and the dimer protein fraction suggesting that they are more likely to be intramolecular. Taken together, these results suggest that C412 is very likely an important residue for the dimerisation process.

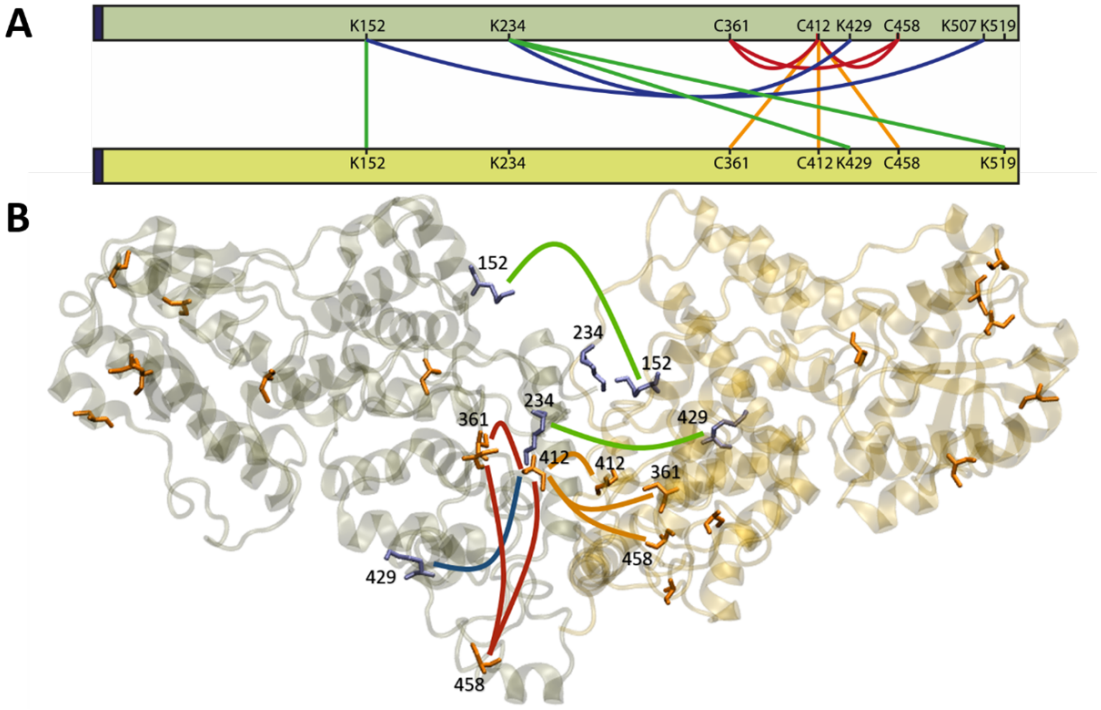

Figure 5.4: Network of cross-links. (**A**) Network plot of cross-linking sites found for *Er*Cry4a. The orange and red lines are potential inter- and intramonomer disulphide bonds, respectively. The inter- and intramonomer cross-links formed by DSBU and DSSO are shown in green and blue, respectively. The dark blue region at the N-terminus (left hand side) corresponds to the five additional amino acids that remained after His-tag cleavage. (**B**) Graphical representation of a possible coordination of two *Er*Cry4a monomers. The cross-links are drawn with the same colour code as in (A). Cysteine residues are shown in orange and lysine residues in purple. Only residues 1-495 are displayed.

5.2.3 Focus on C317 and C412

As C317, the most solvent-exposed cysteine residue of *Er*Cry4a, repeatedly came up as an important potential link between two monomeric units using theoretical approaches and as C412 seemed to be a residue that stands out in experimental approaches, those two residues were chosen for a more detailed investigation by analysing the mutants C412A, C317A, and C412A+C317A. Fig. 5.2C shows that

Table 5.3: XL-MS results. Only disulphide bonds and cross-links between peptides containing lysine residues [25] that were identified with a MeroX [26] score greater than 50 were considered candidates for further analysis. The residues in brackets clarify the position of the linked residue within each peptide.
*A disulphide bond between two C412 residues was found in both the monomer and dimer fractions, possibly due to cross-contamination between the two fractions on the SDS-PAGE gel.

Link	Experiment	MeroX score	Peptide 1	Peptide 2	Linked residues
DSBU	1 (dimer)	81	144-158 (K14)	144-158 (K14)	K152-K152
		76	224-241 (K16)	434-440 (K1)	K234-K429
DSBU	1 (monomer)	59	224-241 (K16)	434-440 (K1)	K234-K429
DSSO	1 (dimer)	52	224-241 (K16)	516-526 (K9)	K234-K519
DSSO	1 (monomer)	109	144-158 (Y13)	503-514 (K10)	K152-K507
disulphide	1 (dimer)	144, 106, 90	415-420 (C3)	415-420 (C3)	C412-C412*
		103, 97	362-370 (C5)	415-420 (C3)	C361-C412
		120	460-467 (C4)	415-420 (C3)	C458-C412
disulphide	1 (monomer)	121, 97, 90, 79	415-420 (C3)	415-420 (C3)	C412-C412*
		117	460-467 (C4)	415-420 (C3)	C458-C412
disulphide	2 (monomer)	145, 101, 97, 94, 93	362-370 (C5)	415-420 (C3)	C361-C412
		164, 149, 81	415-420 (C3)	415-420 (C3)	C412-C412
		166, 109	460-467 (C4)	415-420 (C3)	C458-C412
		95	362-370 (C5)	460-467 (C4)	C361-C458

they all still experience dimerisation and Fig. 5.5 shows that all three mutants experience a lower degree of dimerisation than *Er*Cry4a WT if analysed under exactly the same conditions. Furthermore, the double mutant dimerises less than both single mutants in both MP (Fig. 5.5B) and denaturing gel (Fig. 5.5C) results with the degree of dimer formation following the order WT > C317A > C412A > C412A+C317A for the MP measurements. On the SDS-PAGE gel (Fig. 5.5C) the order of C317A and C412A is reversed and might indicate a different preference for covalently and non-covalently linked dimers to form for the different mutants. The reason for this being that the MP measurements could include non-covalently linked dimers (although only if bound strongly due to the very low protein concentrations of 13-15 nM used for these experiments) whereas a denaturing gel would not. Furthermore, the presence of trimers in the SDS-PAGE gel suggests the involvement of several disulphide links.

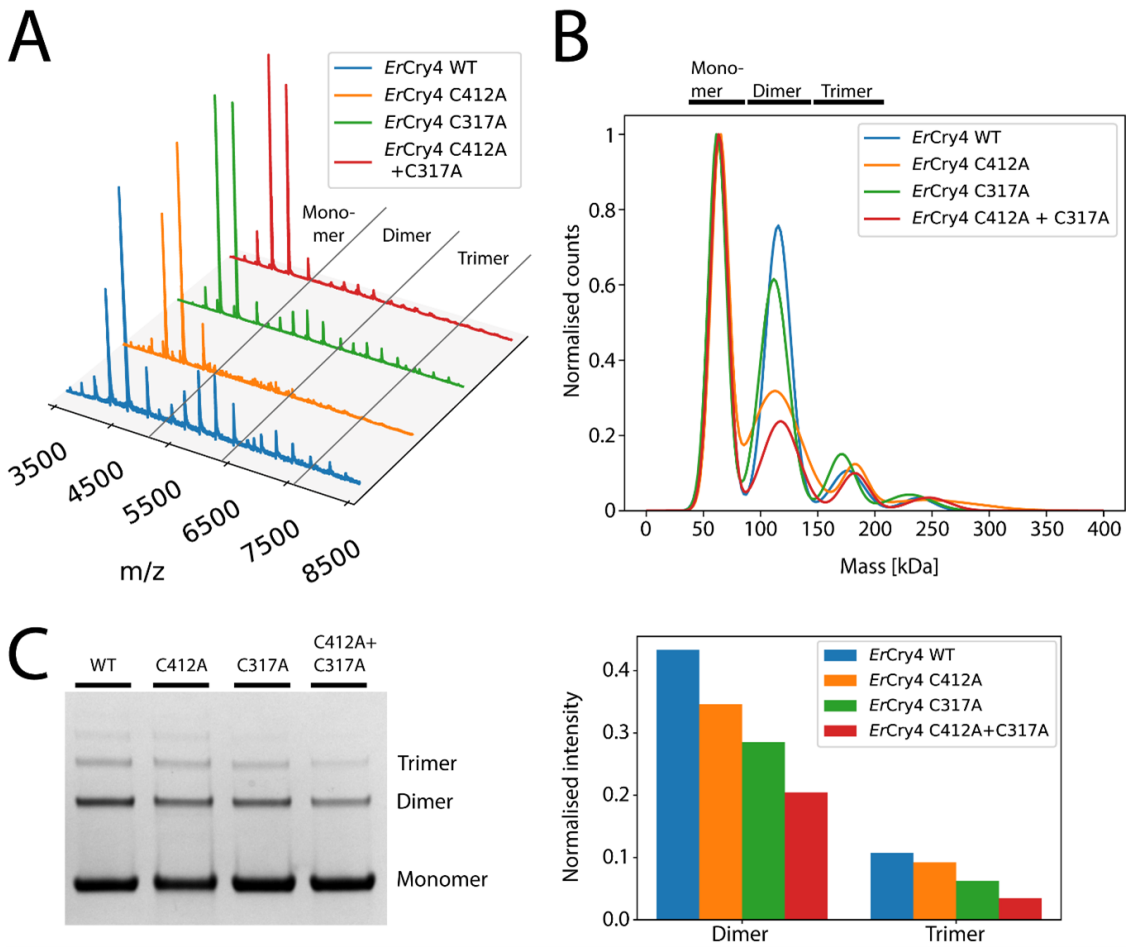

Figure 5.5: Focus on C317 and C412. (**A**) Native mass spectra of WT *Er*Cry4a and three cysteine mutants. All four proteins are mostly monomeric but also show charge-state envelopes from dimers and trimers. (**B**) Mass photometry results for the same proteins at concentrations between 13-15 nM displayed as Gaussian-fitted normalised counts. The four traces have been scaled so that the monomer peaks have the same height. The results have been fitted using a custom python script written by Josh Bishop. (**C**) Denaturing SDS-PAGE gel for the same four proteins (full SDS-PAGE gels can be found in Fig.5.2) and the normalised intensities of the protein areas.

5.2.4 Dimerisation under different light conditions

Interestingly, when investigating all of the cysteine mutants that were expressed during the process of this collaborative work under different light conditions over time, it became evident that all of the mutants mentioned in this work dimerise over time in blue light with the only exception being *Er*Cry4a R356E, a mutant that does not bind the cofactor (henceforth referred to as apo-protein), whereas the dimerisation process is not promoted when the samples are left in the dark.

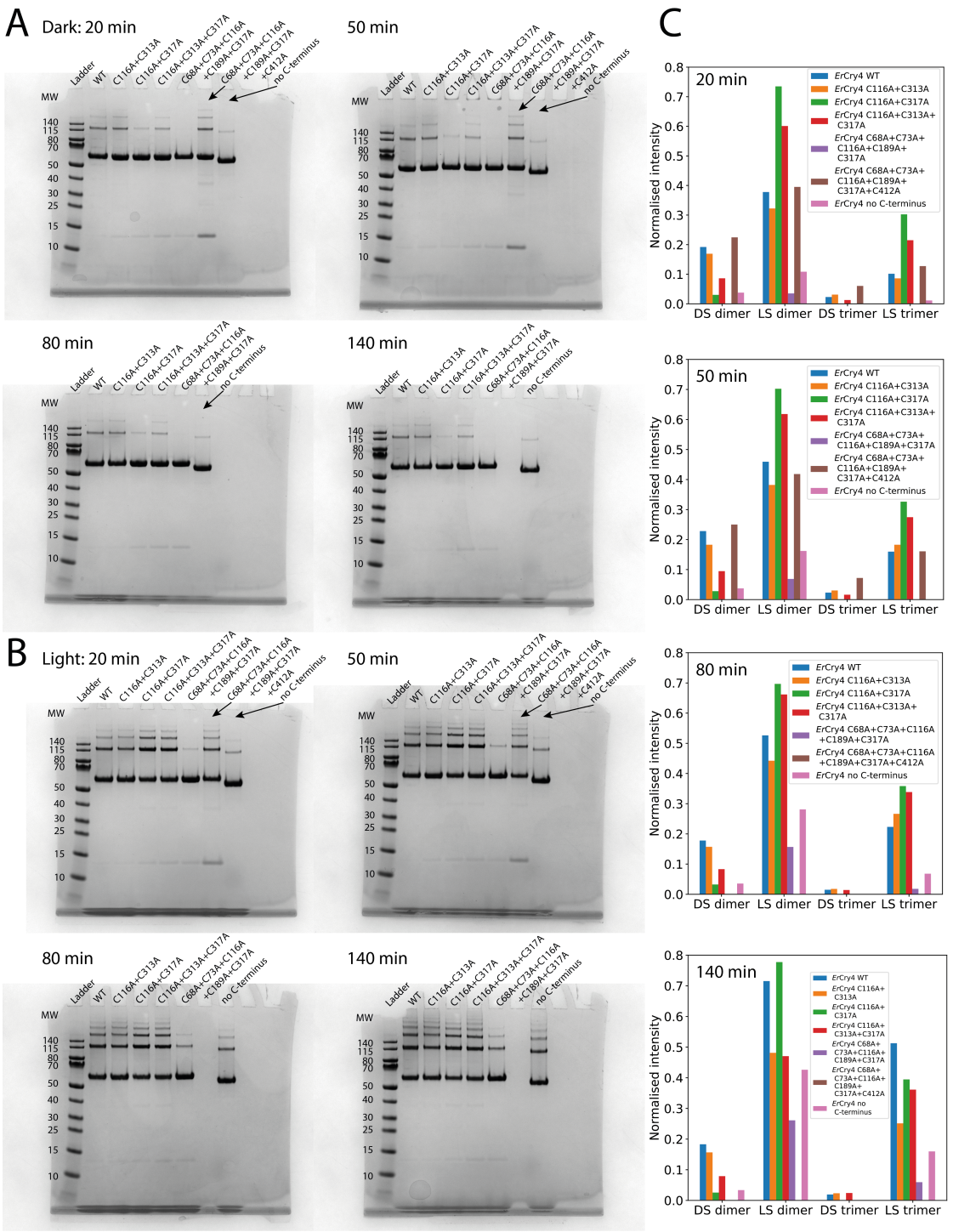

Figure 5.6: Denaturing SDS-PAGE gels monitoring dimerisation over time in the dark and under blue-light conditions of different *Er*Cry4a mutants. (**A**) Dimerisation over time in darkness. (**B**) Dimerisation over time in blue light. For the blue light exposure, a 450 nm LED at about 0.3-1 mW was used. (**C**) Normalised intensities of the protein areas visible in (A) and (B). Each dimer and trimer fraction of a protein was normalised with respect to its monomer fraction on the same SDS-PAGE gel. The SDS-PAGE gel displayed in (A) after an incubation time of 20 min is the full gel as displayed in Fig. 5.2B. The protein indicated as having no C-terminus ends at residue 496.

The results showing this are displayed in Figures 5.6, 5.7, and 5.8.

For the first two figures, the dimerisation of different *Er*Cry4a mutants and the WT over time was monitored in the dark and under blue-light conditions using SDS-PAGE gels with an initial 20 min incubation period in both states. Further SDS-PAGE gels were run for both states at 50 min, 80 min, and 140 min after the start of the incubation time. Fig. 5.6 monitors the WT, the C116A+C313A, C116A+C317A, C116A+C313A+C317A, C68A+C73A+C116A+C189A+C317A (5-Cys), and the C68A+C73A+C116A+C189A+C317A+C412A (6-Cys) mutant, as well as a truncated WT variant which lacks the C-terminal part after residue 496. Fig. 5.7 follows the degree of dimerisation over time of the WT, the C317A, C412A, C116A+C317A, and the 5-Cys mutant. With this Fig. 5.7 is partially a repeat of Fig. 5.6. Both figures also feature a panel C which displays the normalised itensities of the dimer and trimer fractions with respect to the respective monomer fraction for each protein. However, the analysis of small intensity differences of SDS-PAGE gel bands is prone to errors and the results should therefore be taken with a grain of salt. In general, all proteins show dimerisation over time under blue-light conditons with the dark state conditions not being able to promote increased dimerisation. One single exception is visible in Figure 5.7 for the C116A+C317A mutant which does not seem to dimerise over time in blue light. However, this turned out to be most likely due to a mix up of different buffer vials when the samples were prepared for shipping. The proteins that have biological repeats across both figures show that despite the same initial light conditions, there are differences in the initial degree of dimerisation as can be seen comparing the SDS-PAGE gels taken after the inital 20 min of incubation time, e.g. when comparing both *Er*Cry4a WT with the one in Fig. 5.6 showing a higher degree of initial dimerisation as in Fig. 5.7. This points out the importance of having exactly the same initial conditions for quantitative comparisons.

Qualitatively, it is interesting to note that the 6-Cys mutant (where present) shows a comparatively high degree of dimerisation and higher-order oligomerisation with the 5-Cys mutant showing the opposite. It is possible that this means that

the 6-Cys mutant, which has all residues found to be surface exposed (see Fig. 5.3) mutated to alanines, has adopted a slightly different shape and might not be functionally relevant anymore. Notably, the variant lacking the C-terminus seems to dimerise as well but forms higher-order oligomers to a lesser extent than other protein variants (with the exception of the 5-Cys mutant).

Figure 5.8 features all of the mutants mentioned in this work and shows that they all dimerise over time in blue light with the only exception being *Er*Cry4a R356E, the apo-protein. Panel A shows how all proteins are mostly monomeric in the dark, then exhibit dimers and sometimes trimers in Panel C after light exposure, accompanied by the result displayed in panel D showing the irreversibility of the the covalent oligomerisation processes. The longer the light exposure, the more oligomers and higher order oligomers are being formed as panel B points out, at least in the cases of *Er*Cry4a WT and *Cc*Cry4a WT. However, some monomer and dimer are still visible even after 24 hr of light exposure. Most panels of Figure 5.8 seem to suggest that the *Gg*Cry4a proteins dimerise to a larger degree than the other proteins tested. However, this observation needs to be interpreted with caution as all proteins tested had been stored for different amounts of time at the time the SDS-PAGE gel was run and therefore the results for each protein are not directly comparable.

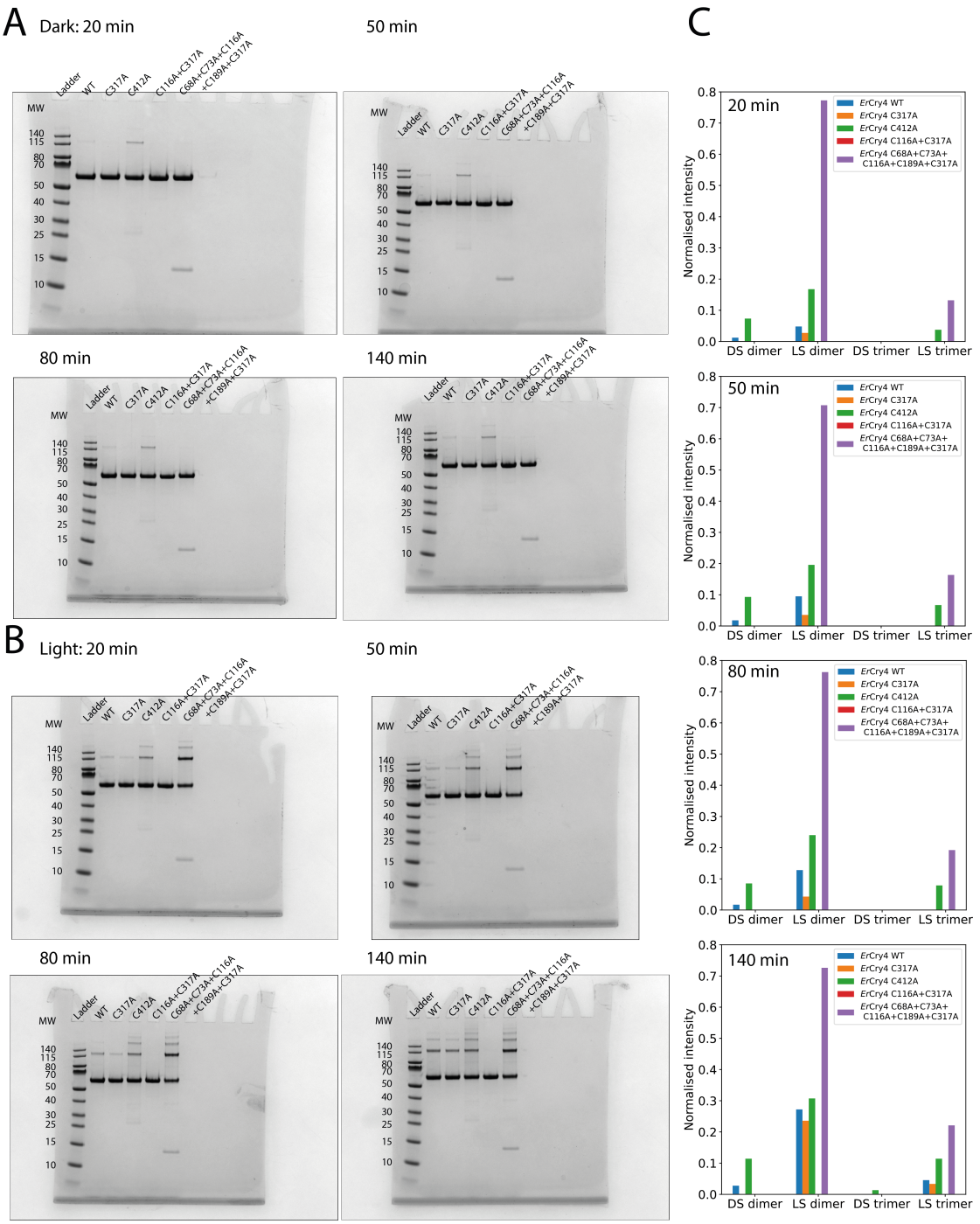

Figure 5.7: Further SDS-PAGE gels monitoring dimerisation over time in the dark and under blue-light conditions of different *Er*Cry4a mutants on denaturing SDS-PAGE gels. (**A**) Dimerisation over time in darkness. (**B**) Dimerisation over time in blue light. For the blue light exposure, a 450 nm LED at about 0.3-1 mW was used. (**C**) Normalised intensities of the protein areas visible in (A) and (B). Each dimer and trimer fraction of a protein was normalised with respect to its monomer fraction on the same SDS-PAGE gel. The SDS-PAGE gel displayed in (A) after an incubation time of 20 min is the same as displayed in Fig. 5.2C.

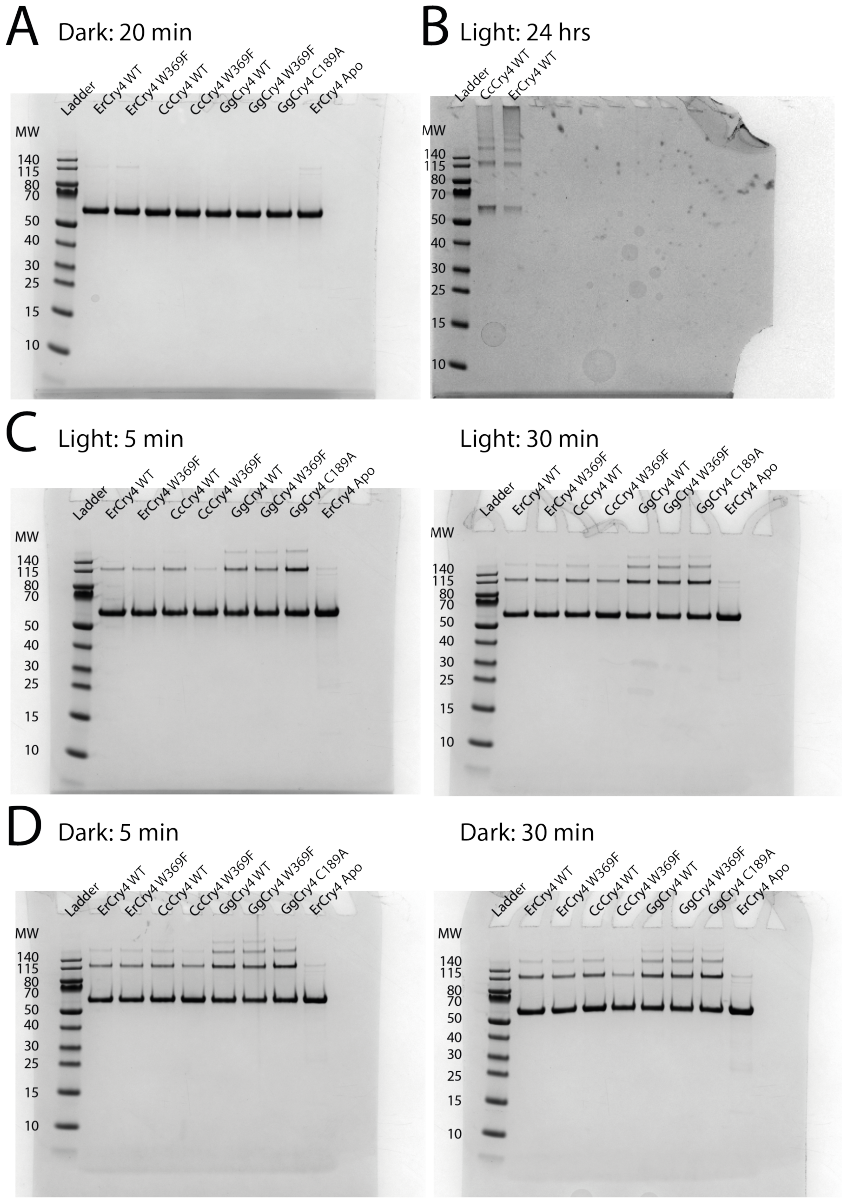

Figure 5.8: Extent of dimerisation of different avian cryptochrome species on denaturing SDS-PAGE gels. (**A**) Proteins after 20 min in darkness before light exposure. (**B**) Light exposure for about 24 hours. Figure courtesy of Laila M. N. Shah. (**C**) Light exposure for 5 and 30 min after an initial 20 min in the dark. (**D**) Samples kept in darkness for 5 and 30 min after 30 min of light exposure. For the light exposure a 450 nm LED at about 0.3-1 mW was used. *Er*Cry4 Apo refers to *Er*Cry4 R356E, a protein lacking the FAD cofactor.

5.2.5 The search for the interaction interface

Electron paramagnetic resonance spectroscopy (EPR), in particular continuous wave EPR(cw-EPR) and double electron-electron resonance (DEER), were used to investigate the distances between paramagnetic, permanent spin centres in monomers and dimers of Cry4a. The original aim was to investigate the interaction surface between two monomer units and to use DEER and HDX-MS as two complementary techniques for this purpose. HDX-MS alone could yield information on the interaction surface, but not about the relative orientation of the two monomers, which would be the question that DEER could help answering. However, due to problems with a size-exclusion column needed on the HDX-MS setup in order to separate the dimer from the monomer, the HDX-MS experiment could not successfully be performed up until now and this section will therefore only focus on the initial EPR results. The realisation of the experiments presented in this section was only possible with the help of Arnau Bertran (a researcher in the groups of C. Tait, A. Bowen and C. Timmel at the University of Oxford) who made it possible that the first DEER experiment on labelled proteins was recorded as well as one on the unlabelled protein. The second DEER experiment on labelled proteins could only be performed thanks to the help of William Myers, the manager of the Centre for Advanced ESR (CAESR) at the University of Oxford within which all experiments were performed.

DEER was performed on the *Er*Cry4a mutant mentioned above where five cysteines had been replaced with alanines (*Er*Cry4a 5-Cys, with each monomer unit comprising the following mutations: C68A, C73A, C116A, C189A, and C317A). This mutant was chosen as it was expected to have C412 left as the only surface-exposed residue, with all other cysteines being buried inside of the protein. Based on the above presented results, dimerisation via a disulphide bond involving C412 would be expected. If C412 is the only surface accessible residue, labelling before dimerisation would lead to the sulphur atom of C412 being unavailable for disulphide bond formation. In this case, it would be expected that DEER would not be able to extract any distance measurements given that dimerisation would be prevented.

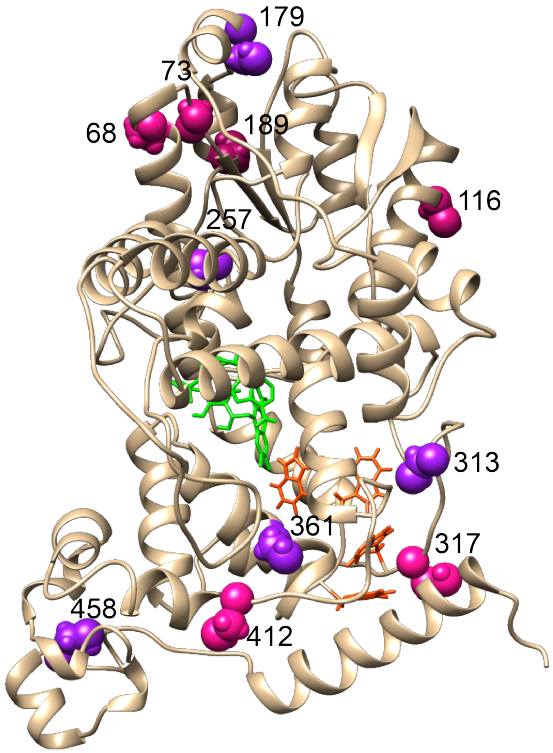

Figure 5.9: Cysteines of *Er*Cry4a WT. Surface-exposed cysteines displayed in pink, buried cysteines displayed in purple, FAD displayed in green and the Trp tetrad in orange. Cysteine exposure defined according to the solvent exposure results of Maja Hanić displayed in Figure 5.3 above. Only residues 1-495 are displayed. The same structure as in Figure 5.3 was used.

If, however, conformational changes initiated upon light-excitation and radical formation allowed other cysteine residues to become surface-exposed, and hence available for dimerisation via disulphide bridge formation, a successful distance measurement would be expected. Another possible outcome would be the formation of non-covalent dimers following light-excitation which would also lead to measurable distances. Figure 5.9 shows where all cysteines, surface accessible and buried, are located. *Er*Cry4a 5-Cys was produced by the collaborating laboratory in Oldenburg and shipped without any reducing agents. The protein was subsequently labelled with the cysteine-conjugating nitroxide radical label (iodoacetamido)proxyl (IPSL) [27] in the dark and the labelling success was assessed using both native MS and continuous wave EPR (cw-EPR). Previously, another nitroxide spin label was trialled as well ((2,2,5,5-tetramethylpyrroline-1-oxyl-3-methyl)- methanethiosulfonate or

MTSL). However, this label consistently lead to over-labelling with three or more spin labels attached to every protein, even at very short labelling incubation periods. The cw-EPR results of the labelled sample indicated that only a fraction of the protein had been successfully labelled (data not shown). However, the native MS results showed that the label was physically attached to almost all protein molecules for the first DEER experiment performed and to a large fraction of the proteins for the second experiment. Figure 5.10A shows these results. Unfortunately, due to a problem with the balance used for weighing out the needed amount of IPSL at the time, the protein for the first EPR experiment was over-labelled as it also contained doubly instead of only singly labelled protein. Panel B of the same figure shows a curious finding observed during the labelling process: before the labelling the amount of dimer present in the sample is a lot higher than afterwards. This could indicate a reason for the discrepancy between the labelling success measured using native MS and cw-EPR. The presence of a spin label radical might reduce the disulphide bonds of the protein and thereby decrease the amount of dimer in the sample.

In preparation for two of the DEER experiments on *Er*Cry4a 5-Cys labelled with IPSL, half of the proteins were kept in complete darkness and the other half was exposed to blue light after labelling to induce dimerisation. A third DEER experiment was performed on an unlabelled sample that had been exposed to blue light.

However, after the initial preparations another problem arose which is displayed in the electron spin-echo field-swept spectrum in Figure 5.11A: the blue spectrum, taken of the protein kept in its DS, has the typical shape expected for a nitroxide label, but the red spectrum shows a distinct and large second peak at around 1209 mT which might be caused by $FADH^{\bullet}$ as suggested by a comparison with a previously recorded spectrum of this component by Tommy L. Pitcher in the Appendix Figure A.40. This problem arose for the first DEER experiment performed and was greatly reduced for the second attempt by letting the light state (LS) sample rest in darkness overnight before measurement (see Appendix Figure A.41). Figure 5.11B indicates the intensities of the pump spin ensemble measured during the sample characterisation stages of the labelled dark state (DS + IPSL), the labelled

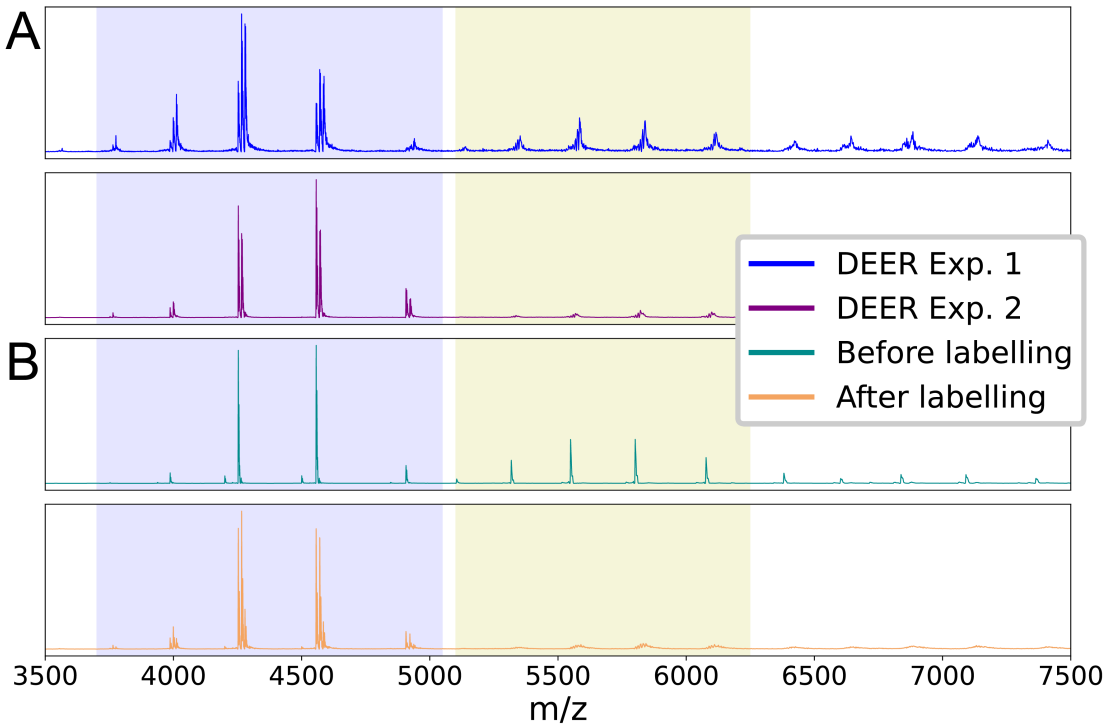

Figure 5.10: Native MS of IPSL-labelled *Er*Cry4a 5-Cys. Theoretical mass of *Er*Cry4a 5-Cys: 63909.47 Da. Theoretical additional mass of attached label: 198.25 Da. (**A**) Labelled samples as used for 4-pulse DEER experiments. The sample used for the first experiment contained singly labelled protein (64005.86±26.65 Da), but also approximately the same amount of doubly labelled protein (64198.18±19.28 Da) and a small fraction of leftover unlabelled protein (63787.45±16.88 Da). The sample for the second experiment contained mostly unlabelled protein (63780.09±2.23 Da), some singly labelled protein (about two thirds the amount of unlabelled sample) (63999.57±19.06 Da), and no doubly labelled protein. (**B**) *Er*Cry4a 5-Cys at different stages during the labelling process: before (upper panel) and after (lower panel) labelling with IPSL. The two panels show the contrasting degree of dimerisation which is drastically reduced upon addition of IPSL.

LS (LS + IPSL) and the unlabelled LS (here termed LS) to confirm the presence of a spin signal before DEER-measurement.

The results of the first attempt displayed in Figure 5.12 show what was unfortunately to be expected: due to the doubly labelled proteins both the LS and DS results show multiple distances and it is difficult to distinguish which distance might be located with which residues within the protein. An additional difficulty is posed by the supposed presence of $FADH^{\bullet}$. Table 5.4 makes a first attempt at elucidating possible distances that might have been measured. The

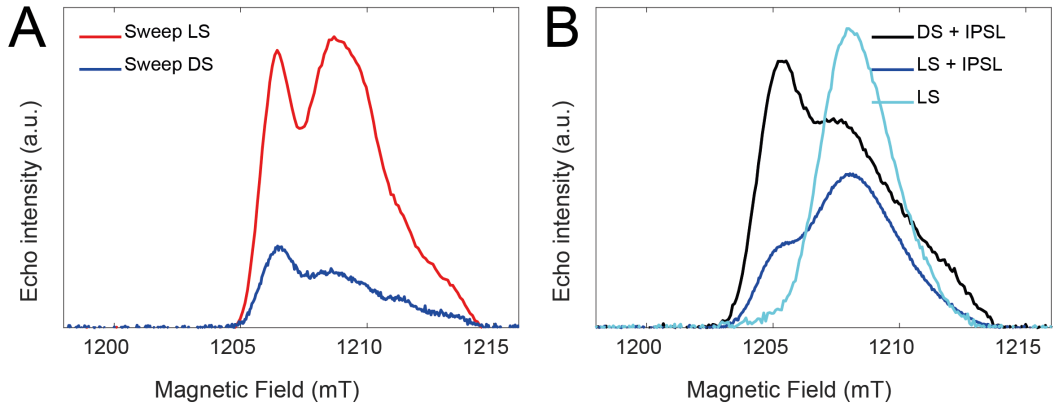

Figure 5.11: Electron spin-echo field-swept spectra. (**A**) Two magnetic field sweeps for the light state (LS) and the dark state (DS) taken while performing the first experiment on samples labelled with IPSL. (**B**) Intensities of the pump spin ensemble to confirm the presence of a spin signal, here shown for the light state (LS) and dark state (DS) of the first experiment performed on samples labelled with IPSL. In light blue it features a trace taken for an unlabelled LS sample.

table only shows intra-molecular distances based on the homology model displayed in Figures 5.3 and 5.9 as opposed to inter-molecular distances which would be of higher interest and is therefore at its current stage merely a playground for ideas. Possible distances that could be hiding in the data set are the distances between the FAD cofactor and either C257 or C313, as these are both approximately 17 Å, corresponding to the best resolved peak of the LS distance distribution in Figure 5.12. The distances between C412 and C313, C458, or the cofactor could all be hiding in the second area of possible distances identified within the same figure. However, this analysis is very approximate, not least since the table also only gives a rough estimate based on distances of only one frame of a homology model in the dark state, not considering any potential conformational changes as well as only considering specific atoms as the approximate spin centres.

To elucidate the distance between two FAD spin centres within a dimer, unlabelled *Er*Cry4a 5-Cys was subjected to blue light and subsequently measured. The results displayed in Figure 5.13 show that there was no discernible distance to be extracted from the data. However, given that there was a spin signal corresponding to the $FADH^{\bullet}$ radical (as displayed in Figure 5.11B), this leads to the assumption

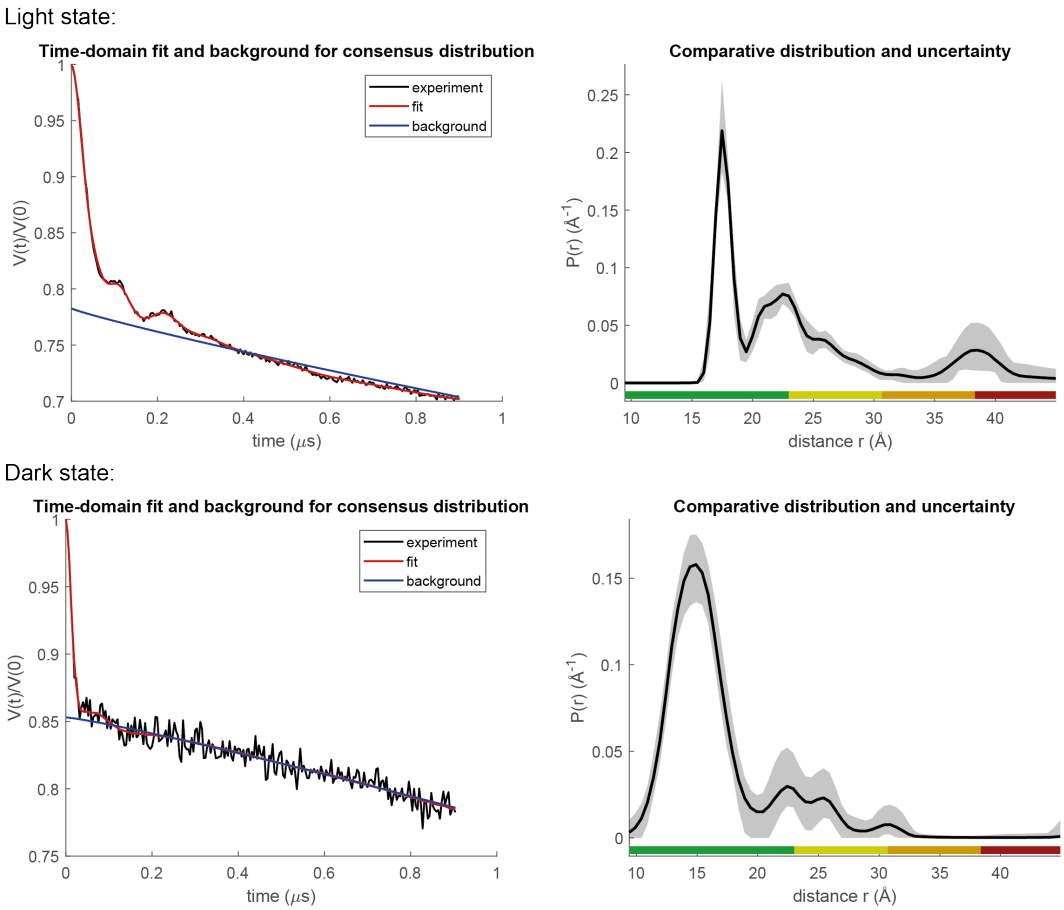

Figure 5.12: Distances obtained for *Er*Cry4a 5-Cys using DEER (first experiment). DEER traces for the light and dark state protein results as indicated on the left. Distance distributions and uncertainties are displayed on the right and calculated using Tikhonov regularisation as well as DEERNet Spinach, a method based on pre-trained neural networks, in DeerAnalysis [28]. The plots show the joined results of both methods. The colour coding below the distance distribution plots indicates the certainty with which the distances can be determined and the grey area around the black curve indicates the uncertainty of the finding.

that an *Er*Cry4a 5-Cys dimer generally only contains one FAD in its radical form or none at all. However, given the spin signal in Figure 5.11B, the latter is rather unlikely. In the latter case, the $FADH^{\bullet}$ radical signal in Figure 5.11B would then be attributed to semi-reduced flavin in the monomers only.

The second attempt to measure distances using DEER, displayed in Figure 5.14, shows fewer and more distinct distances measured for the LS at approximately 17 Å and 22.5 Å and the absence of a (although uncertain) distance measured at around 38 Å, which could potentially indicate the absence of a dipolar interaction

Table 5.4: Distances between Cα atoms and between one nitrogen of the isoalloxazine moiety of FAD (N10, defined as one of the N in the central isoalloxazine ring [29]) and the Cα atoms of all available cysteine residues within *Er*Cry4a 5-Cys. The distances were measured from the homology model displayed in Figure 5.9 and are therefore only approximate measures of the actual distances within the protein.

Residue 1	Residue 2	Distance / Å
412 (Cα)	179 (Cα)	59.6
412 (Cα)	257 (Cα)	37.6
412 (Cα)	313 (Cα)	25.6
412 (Cα)	361 (Cα)	10.1
412 (Cα)	458 (Cα)	21.3
FAD (N10)	179 (Cα)	37.9
FAD (N10)	257 (Cα)	17.1
FAD (N10)	313 (Cα)	17.4
FAD (N10)	361 (Cα)	14.3
FAD (N10)	412 (Cα)	21.8
FAD (N10)	458 (Cα)	31.4

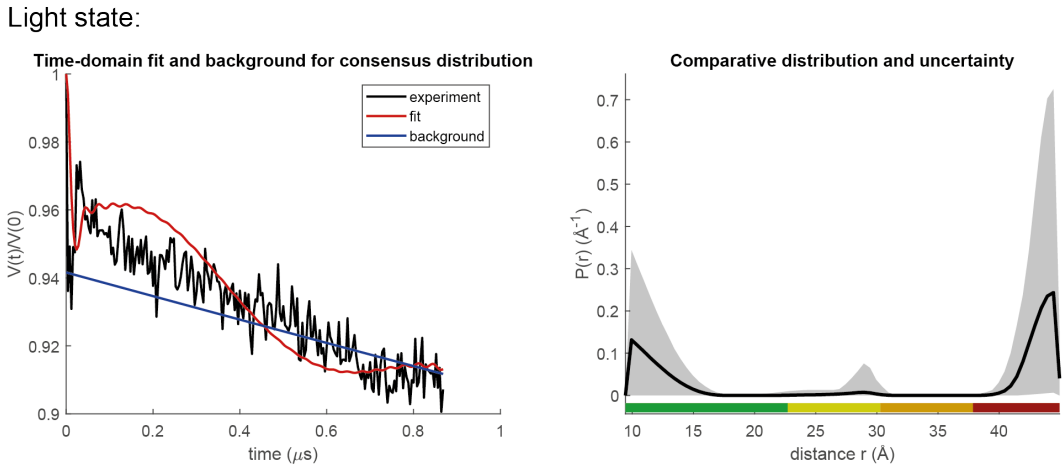

Figure 5.13: Distances obtained for unlabelled *Er*Cry4a 5-Cys using DEER. DEER traces obtained for the light state protein. Distance distributions and uncertainties are displayed on the right and calculated using DeerAnalysis [28]. For further details see Figure 5.12. For this data, a manual fit using only Tikhonov Regularization within the Matlab® *DeerAnalysis2022* routine was applied. However, this analysis can be user-biased and did not lead to a any discernible distance distributions either. Most likely, the difficulties in finding a fit show that there is no clear distance to fit to.

between two FAD cofactor radicals. In contrast, the DS results in Figure 5.14 show no distances that can be detected with certainty whereas in the first experimental attempt multiple distances were extracted from the data. Firstly, this shows that dimerisation was achieved in the LS whereas in the DS no two radicals seem to have been sufficiently close to one another to interact. Secondly, it hints at the possibility to have several orientations within two monomer units that form a dimer as not only one but two distinct distances were found in the LS.

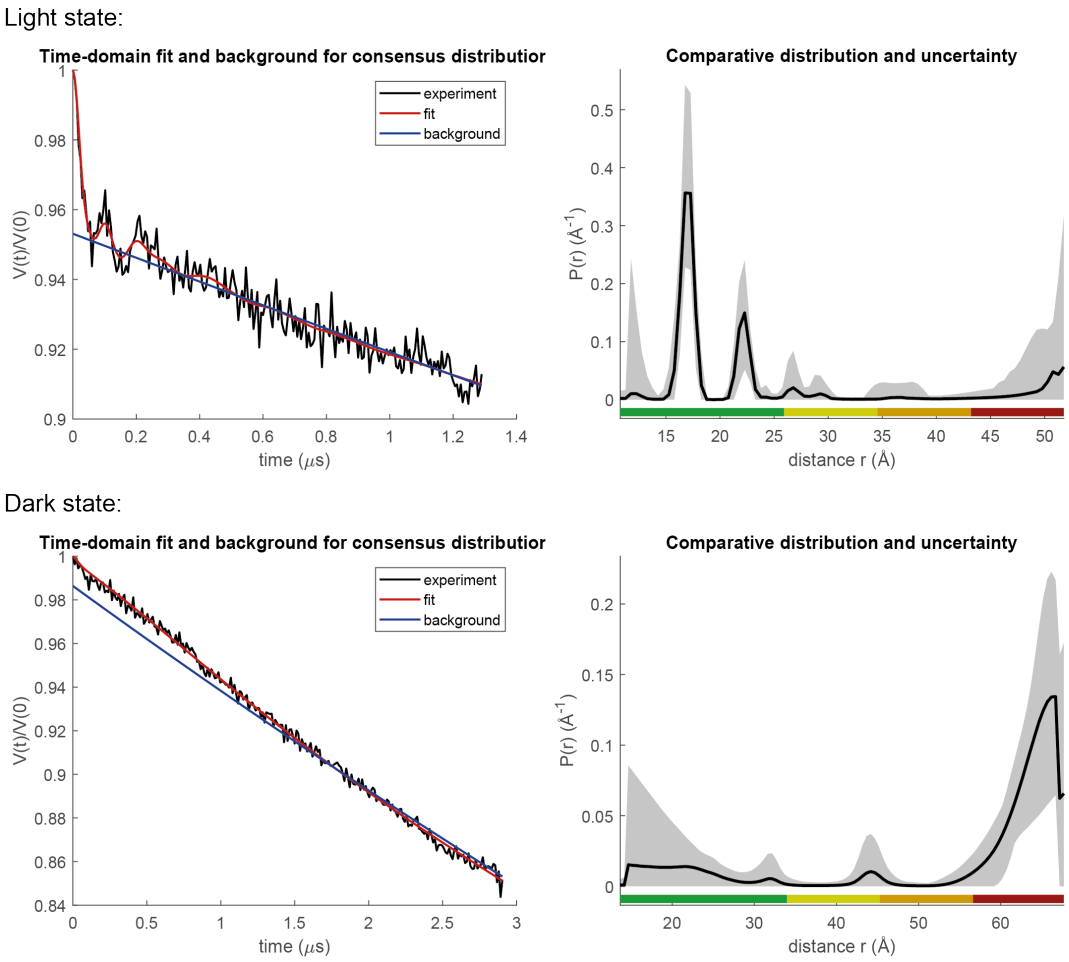

Figure 5.14: Distances obtained for *Er*Cry4a 5-Cys using DEER (second experiment). DEER traces for the light and dark state protein results as indicated on the left. Distance distributions and uncertainties are displayed on the right and calculated using DeerAnalysis [28]. For further details see Figure 5.12.

Summing up the findings of these preliminary results aiming to find the interaction interface of the *Er*Cry4a 5-Cys dimers, it is clear from the data that dimerisation induced by simple blue light exposure is sufficient for DEER to detect

signals that are not present in the DS of the protein and that these distances can be used in conjunction with experimental results from other techniques (such as HDX-MS) to find out which residues are approximately 17 Å and 22.5 Å from each other. A strong candidate would be C412. It could be entirely possible that C412 is indeed the only labelled residue and two C412 could be approximately 17 Å and 22.5 Å from each other within two different dimeric conformations. Also important to consider is the fact that sometimes trimers and higher order oligomers are formed as well, although to a lesser extent. Another possibility is that *Er*Cry4a 5-Cys might form chains with two different interfaces to two neighbours which could also explain two different distances in a sample where mostly C412 is expected to be labelled. However, this work requires further analysis including an analysis of the extent of dimerisation and oligomerisation within the samples used. The protein containing only one labelled residue was planned to be digested on the HDX-MS setup to search its peptides for the position of the spin label. However, due to frequent breakdowns of the instrument it was so far not possible to finish this analysis successfully. Subsequently, a simulation of attached nitroxide label separations within an *Er*Cry4a 5-Cys dimer would be useful. For this, a generation of possible dimer structures using AlphaFold [30] together with the results presented above might further assist the interpretation of the results. Lastly, including other proteins in this analysis might be worth a consideration such as pigeon Cry4a which, out of a comparison with chicken, quail, and robin Cry4a proteins, is the only protein that has only C412 but no other cysteine residues on its surface (assuming that all these proteins have a very similar conformational structure). It could be trialled for nitroxide labelling given no complicated cysteine mutant would be necessary to achieve the same results as for *Er*Cry4a 5-Cys. However, this protein has another cysteine residue towards the C-terminal tail that could potentially cause unexpected trouble. All in all, these initial results are promising and represent a first small step towards finding the interaction surface of two *Er*Cry4a 5-Cys monomers.

5.3 Conclusion

The experimental results presented in this chapter show that *Er*Cry4a forms dimers and that to a large extent they are covalently linked via disulphide bonds. However, the results do not exclude the possibility of the presence of non-covalently linked dimers. Theoretical approaches suggest that non-covalently linked dimers would have the potential to be relatively stable, at least in the conformations found and listed in Table 5.2. Another interesting finding is that all proteins and protein variants investigated in this chapter dimerise with the exception of *Er*Cry4 R356E, the apo-protein lacking the cofactor FAD suggesting that excitation of the cofactor induces dimerisation in blue light.

Several cysteines seem to be involved in the formation of disulphide bridges, among which C317 and C412 seem to be important residues as their removal reduces the extent of dimerisation in experiments and their presence promotes stable dimeric units in theoretical approaches. However, the finding of important residues could not answer the question of which other residues might also be involved in dimerisation. Additionally, higher-order oligomers such as trimers were detected which leads to the conclusion that the protein might have more than one interaction surface and possibly more than two cysteine residues involved in dimerisation. The question as to which extent these dimers (or higher-order oligomers) are functionally relevant and important with respect to magnetoreception remains unanswered.

Since dimers do not seem to be easily disrupted in reducing environments (Fig. 5.1C) and as there are other examples of biologically significant intra-cellular disulphide-linked proteins involved in redox processes despite the generally reducing environment of the cell [31–33], it is entirely possible that *Er*Cry4a oligomerisation might have a biological function. However, as can be seen in all the figures displaying a comparison of monomer, dimer and potentially higher order oligomers, only a proportion of the protein seems to dimerise under the experimental conditions chosen. Since the MS charge state envelopes of the dimers look similar to those of the monomers and do not display higher charge states, which would be indicative of an unfolded protein, it seems that dimerisation does not alter the protein conformation

significantly and therefore excludes the option of non-specific aggregation. A possibility would be the existence of a monomer-dimer equilibrium that could have a regulatory function in a downstream signalling process. Examples for this would be the homo-oligomerisation of plant cryptochromes [34–38] and the formation of disulphide bonds between *Mm*Cry1 and Period, a protein involved in circadian function, to form a heterodimer [39].

Xu *et al.* [2] found in their Fig. S5 that samples of *Er*Cry4a which were purified with 10 mM BME to keep the samples as monomeric as possible during purification (as was the case for the proteins investigated for the work in this chapter), had weak signals of photo-excited singlet and triplet states (^{S}FAD* and ^{T}FAD*, respectively) alongside the expected radical pair formation while the same proteins purified without BME showed clear evidence of the formation of relatively long-lived excited states on a nanosecond to microsecond timescale. It was suggested that subtle differences in protein conformation when present in its monomeric or oligomeric forms could prevent electron transfer from the tryptophan tetrad to the excited cofactor.

Summing up, the results presented in this chapter suggest that dimerisation of *Er*Cry4a might play a functional role *in vivo* which could have further effects on the interaction with putative downstream interaction partners during the process of magnetoreception. It also presents an incentive for further work in this area as several questions remain open such as what are the exact conditions in a living cell that contains *Er*Cry4a, whether dimerisation and higher order oligomerisation can be seen *in vivo*, if the initial dimerisation process might be a nucleation process to introduce higher-order oligomerisation, what the protein interaction surfaces are, or what the impact of dimerised proteins would be on the longevity of the different possible radicals and the rates of electron transfer, especially during an interaction with potential downstream signalling partners.

5.4 Materials and Methods

5.4.1 Materials

All chemicals and reagents were purchased from Sigma Aldrich (Gilligham, UK). All proteins were expressed and purified in the laboratory of the collaborating research group of Henrik Mouritsen at the Carl von Ossietzky Universität Oldenburg in Germany. Hanić *et al.* [1] details the protein expression and purification methods.

5.4.2 Native MS

The *Er*Cry4a samples used for these experiments were shipped overnight to Oxford on dry ice from the production laboratory at the University of Oldenburg, in a buffer containing 20 mM Tris, 250 mM NaCl, and 20% glycerol at pH 8 unless otherwise indicated. The samples were kept at -80°C until they were thawed on ice and buffer-exchanged into 200 mM ammonium acetate, pH 8, for the native MS measurements using Zeba^{TM} Micro Spin desalting columns with a molecular weight exclusion limit of 40 kDa (ThermoFisher Scientific). All samples included an N-terminal 10×His-tag used for purification of the proteins.

Measurements were performed in Oxford using in-house gold-plated capillaries on a Q Exactiv^{TM} mass spectrometer in positive ion mode. The samples were sprayed using nanoflow electrospray with a source temperature of 150°C and a capillary voltage of 1.0 kV [40]. The higher-energy C-trap dissociation (HCD) cell voltage was 5 V, in-source trapping was set to -200 V to help with the dissociation of small ion adducts, and the noise threshold parameter was 3. Ion transfer optics and voltage gradients throughout the instruments were adjusted for optimum protein transmission.

The mass spectra were processed using the MassLynx V4.2 software (Waters, Wilmslow, UK).

5.4.3 Mass photometry

All samples were shipped as described above except for the *Er*Cry4a WT sample used in Fig. 5.1 which was shipped with the addition of 2 mM DTT to limit higher order oligomerisation during shipping. This sample was diluted to a concentration of ∼44 nM in 200 mM ammonium acetate, pH 8.3, prior to measurement. All other samples were buffer-exchanged into 200 mM ammonium acetate, pH 8.0, and diluted to 13-15 nM. MP measurements were performed essentially as described previously (Young et al [41]), on a ONE^{MP} instrument (Refeyn Ltd, Oxford) in the case of the *Er*Cry4a WT sample in 5.1 and on a TWO^{MP} instrument (Refeyn Ltd, Oxford) in all other cases. Briefly, a borosilicate microscope coverslip (Thorlabs) was cleaned by ultra-sonication for 5 min in a 1:1 mix of ultrapure water and isopropanol followed by another 5 min of ultra-sonication in ultrapure water and then dried under a stream of nitrogen. Cleaned coverslips and Grace Bio-Labs reusable CultureWell^{TM} gaskets (Bio-Labs) were assembled into flow chambers. 10-20 µL of sample were added to the flow chamber and images of a 3.5 × 12.2 µm^2 region of the glass coverslip surface were acquired at 1000 frames s^{-1}. The mass distributions were obtained from contrasts of ratiometric data by calibration using DiscoverMP (Refeyn Ltd, Oxford).

5.4.4 SDS-PAGE

For SDS-PAGE gels, the samples were covered in the dark or subjected to blue light (450 nm, 0.3 mW) for 20 min or longer (if indicated) at room temperature at a concentration of ∼8 µM. 3.2 µg of each sample were then mixed with diluted NuPAGE LDS sample buffer (4X) (Invitrogen, Waltham, MA, USA) before being loaded on NuPAGE Novex 4-12% bis-Tris gels, 1.0 mm x 10 well (Invitrogen) and run with NuPAGE MES SDS running buffer (20X) (Invitrogen). Samples of proteins without a His-tag, containing 2 mM DTT (Sigma-Aldrich) were shipped to Oxford and kept in the light while being incubated with different concentrations of DTT for 20 min at room temperature prior to being submitted to the SDS-PAGE gel. Protein bands were visualised by Coomassie blue staining using QuickBlue Protein Stain (LuBio, Zürich, Switzerland) and accompanied by either a SeeBlue^{TM}

Plus2 Pre-stained Protein Ladder (Invitrogen) or a PageRuler Prestained Protein Ladder (Thermo Scientific, Waltham, MA, USA). The software GelAnalyzer 19.1 [42] was used for evaluation of the gel bands.

5.4.5 Chemical cross-linking and mass spectrometric analysis

Chemical cross-linking followed by mass spectrometric analysis (XL-MS) was used to probe the interaction surfaces of the linked monomers. Samples for XL-MS were sent to Oxford as described above, with the addition of 2 mM DTT in the shipping buffer to limit higher-order oligomerisation. Furthermore, His-tags were cut leaving only 5 additional amino acids at the N-terminus (GAMGS).

To investigate potential disulphide bonds, protein solutions were incubated with the linkers disuccinimidyl sulphoxide (DSSO) [23] or disuccinimidyl dibutyric acid (DSBU) [24] to cross-link primary amines at the dimer interface using the manufacturer's recommended procedure, as described previously [43]. Cross-linked and non-cross-linked proteins were resolved using SDS-PAGE for analysis of the cross-linkers and direct analysis of the disulphide bonds.

Bands were excised from gels and prepared for mass spectrometry as described by Shevchenko *et al.* [44]. After in-gel digestion, peptides were resolubilised in buffer A (H_2O, 0.1% formic acid), and ~5-10 µL of peptide was loaded onto a reverse-phase trap column (Acclaim PepMap 100, 75 µm × 2 cm, NanoViper, C18, 3 µm, 100 Å, ThermoFisher, Waltham, MA, USA) using an Ultimate 3000 autosampler and washed with 40 µL of buffer A at a flow rate of 20 µL min^{-1}. The trapped peptides were then separated by applying a 15 cm reverse-phase analytical column (Acclaim PepMap 100 C18, 3 µm, 150 mm × 0.075 mm) using a 105 min linear gradient from 5% to 55% buffer B (80% acetonitrile, 20% water, 0.1% formic acid) at a flow rate of 300 nL min^{-1}. Long and hydrophobic peptides were washed from the column using 99% buffer B for 15 min at the same flow rate. The separated peptides were electrosprayed in the positive ion mode into an Orbitrap Eclipse Tribrid mass spectrometer (ThermoFisher) for analysis with the help of Tarick J. El-Baba, a

Newton Fellow at the University of Oxford working in the research group of Carol V. Robinson. Fragment ions corresponding to peptides cross-linked by DSBU or DSSO, respectively, were subjected to further sequencing. Further instrumental parameters are detailed in Hanić *et al.* [1]. The cross-linking data were analysed using the free software tool MeroX [26] available at www.StavroX.com.

5.4.6 DEER

To label proteins with nitroxides, 3-(2-Iodoacetamido)-2,2,5,5-tetramethyl-1- pyrrolidinyloxy (IPSL; Sigma-Aldrich) was dissolved in LC-MS grade acetonitrile (ACN) to a concentration of 20 mM and mixed in 10× excess with the protein under dim red light conditions. It was kept on ice in the dark for 1.5 - 2 hr. The excess label was subsequently washed out using Zeba^{TM} Micro Spin desalting columns with a molecular weight exclusion limit of 40 kDa (ThermoFisher Scientific) and the samples were concentrated to above 100 µM.

Prior to DEER measurements the labelling efficiency was determined using an EMX-Micro cw-ESR spectrometer (Bruker) at X-band at room temperature with a critically coupled resonator. Sample handling was performed in the dark or under dim red light conditions. Samples were snap-frozen in liquid nitrogen and kept at -80°C until DEER measurements were performed.

4-pulse DEER [45] was performed using a pulsed spectrometer (ElexSys E580, Bruker) at Q-band and a temperature of 80 K. An Oxford instruments cryostat and an overcoupled dielectric resonantor (D2 EN5107D2, Bruker) were used. DEER traces were acquired using a 4-pulse refocused-echo experiment with the sequence $\frac{\pi}{2} - \tau_1 - \pi - \tau_1 - \pi_{\text{pump}} - \tau_2 - \pi - \tau_2 -$ echo.

For the first DEER experiment, the π detection pulse had a length of 24 ns, the π pump pulse had a length of 12 ns, delays of $\tau_1 = 200$ ns and $\tau_2 = 1000$ ns were used, the external magnetic field was at 12055 G for the LS and at 12054 G for the DS, and the pump pulse was moved in steps of 4 ns with respect to the detection pulse. For the second DEER experiment, the π detection pulse had a length of 14 ns, the π pump pulse had a length of 12 ns, delays of $\tau_1 = 260$ ns and $\tau_2 = 1400$ ns for the LS

and $\tau_2 = 3000$ ns for the DS were used, the external magnetic field was at 12069 G for the LS and at 12050 G for the DS, and the pump pulse was moved in steps of 6 ns with respect to the detection pulse for the LS and in steps of 12 ns for the DS. For the DEER experiment of the unlabelled LS sample, the π detection pulse had a length of 24 ns, the π pump pulse had a length of 12 ns, delays of $\tau_1 = 200$ ns and $\tau_2 = 1000$ ns were used, the external magnetic field was at 12078 G, and the pump pulse was moved in steps of 4 ns with respect to the detection pulse. The LS samples were kept for 1 hr at 450 nm light of ∼0.3-0.7 mW cm^{-2} with the exception of the second DEER LS sample which was kept under these conditions for 2 hrs.

The raw traces were phase- and background-corrected and subsequently analysed using Tikhonov Regularization and DEERNet Spinach using the Matlab® *DeerAnalysis2022* routine to extract the corresponding distance distributions.

References

1. Hanić, M. *et al.* Dimerization of European Robin Cryptochrome 4a. *The Journal of Physical Chemistry B* **127,** 6251–6264. https://pubs.acs.org/doi/10.1021/acs.jpcb.3c01305 (2023).
2. Xu, J. *et al.* Magnetic sensitivity of cryptochrome 4 from a migratory songbird. *Nature* **594,** 535–540. http://dx.doi.org/10.1038/s41586-021-03618-9 (2021).
3. Pepys, M. B. *et al.* Targeting C-reactive protein for the treatment of cardiovascular disease. *Nature* **440,** 1217–1221. http://www.nature.com/articles/nature04672 (2006).
4. Ali, S., Hassan, M., Islam, A. & Ahmad, F. A Review of Methods Available to Estimate Solvent-Accessible Surface Areas of Soluble Proteins in the Folded and Unfolded States. *Current Protein & Peptide Science* **15,** 456–476. http://www.eurekaselect.com/openurl/content.php?genre=article%5C&issn=1389-2037&volume=15%5C&issue=5%5C&spage=456 (2014).
5. Fraczkiewicz, R. & Braun, W. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *Journal of Computational Chemistry* **19,** 319–333. https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1096-987X(199802)19:3%5C%3C319::AID-JCC6%5C%3E3.0.CO;2-W (1998).
6. Schuhmann, F., Korol, V. & Solov'yov, I. A. Introducing Pep McConst—A user-friendly peptide modeler for biophysical applications. *Journal of Computational Chemistry* **42,** 572–580 (2021).

7. Hanić, M., Frederiksen, A., Schuhmann, F. & Solov'yov, I. A. On the energetic differences of avian cryptochromes 4 from selected species. *The European Physical Journal D* **76,** 198. https://link.springer.com/10.1140/epjd/s10053-022-00520-5 (2022).

8. Chen, R. & Weng, Z. A novel shape complementarity scoring function for protein-protein docking. *Proteins: Structure, Function, and Genetics* **51,** 397–408. https://onlinelibrary.wiley.com/doi/10.1002/prot.10334 (2003).

9. Pierce, B. G., Hourai, Y. & Weng, Z. Accelerating Protein Docking in ZDOCK Using an Advanced 3D Convolution Library. *PLoS ONE* **6** (ed Keskin, O.) e24657. https://dx.plos.org/10.1371/journal.pone.0024657 (2011).

10. Pierce, B. G. *et al.* ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. *Bioinformatics* **30,** 1771–1773. https://academic.oup.com/bioinformatics/article/30/12/1771/381628 (2014).

11. Pierce, B., Tong, W. & Weng, Z. M-ZDOCK: a grid-based approach for Cn symmetric multimer docking. *Bioinformatics* **21,** 1472–1478. https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/bti229 (2005).

12. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **26,** 1781–1802 (2005).

13. Phillips, J. C. *et al.* Scalable molecular dynamics on CPU and GPU architectures with NAMD. *The Journal of Chemical Physics* **153,** 044130. https://doi.org/10.1063/5.0014475%20http://aip.scitation.org/doi/10.1063/5.0014475 (2020).

14. Zoltowski, B. D. *et al.* Chemical and structural analysis of a photoactive vertebrate cryptochrome from pigeon. *Proceedings of the National Academy of Sciences* **116,** 19449–19457. http://www.pnas.org/lookup/doi/10.1073/pnas.1907875116 (2019).

15. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9,** 40. https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-9-40 (2008).

16. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* **5,** 725–738. http://www.nature.com/articles/nprot.2010.5 (2010).

17. Yang, J. *et al.* The I-TASSER Suite: protein structure and function prediction. *Nature Methods* **12,** 7–8. http://www.nature.com/articles/nmeth.3213 (2015).

18. Yang, J. & Zhang, Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Research* **43,** W174–W181. https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkv342 (2015).

19. Miller, S. *et al.* Isoform-selective regulation of mammalian cryptochromes. *Nature Chemical Biology* **16,** 676–685. http://dx.doi.org/10.1038/s41589-020-0505-1%20http://www.nature.com/articles/s41589-020-0505-1 (2020).

20. *Protein dimerization and oligomerization in biology* (ed Matthews, J. M.) (Springer Science & Business Media, 2012).

21. Sousa, S. F., Neves, R. P., Waheed, S. O., Fernandes, P. A. & Ramos, M. J. Structural and mechanistic aspects of S-S bonds in the thioredoxin-like family of proteins. *Biological Chemistry* **400,** 575–587. https://www.degruyter.com/document/doi/10.1515/hsz-2018-0319/html (2019).

22. Kabsch, W. A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography* **34,** 827–828 (1978).

23. Kao, A. *et al.* Development of a Novel Cross-linking Strategy for Fast and Accurate Identification of Cross-linked Peptides of Protein Complexes. *Molecular & Cellular Proteomics* **10,** M110.002170. http://dx.doi.org/10.1074/mcp.M110.002212%20https://linkinghub.elsevier.com/retrieve/pii/S1535947620313724 (2011).

24. Müller, M. Q., Dreiocker, F., Ihling, C. H., Schäfer, M. & Sinz, A. Cleavable Cross-Linker for Protein Structure Analysis: Reliable Identification of Cross-Linking Products by Tandem MS. *Analytical Chemistry* **82,** 6958–6968. https://pubs.acs.org/doi/10.1021/ac101241t (2010).

25. Mädler, S., Bich, C., Touboul, D. & Zenobi, R. Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *Journal of Mass Spectrometry* **44,** 694–706. https://onlinelibrary.wiley.com/doi/10.1002/jms.1544 (2009).

26. Götze, M. *et al.* Automated Assignment of MS/MS Cleavable Cross-Links in Protein 3D-Structure Analysis. *Journal of the American Society for Mass Spectrometry* **26,** 83–97. https://pubs.acs.org/doi/10.1007/s13361-014-1001-1 (2015).

27. Ackermann, K., Chapman, A. & Bode, B. E. A Comparison of Cysteine-Conjugated Nitroxide Spin Labels for Pulse Dipolar EPR Spectroscopy. *Molecules* **26,** 7534. https://www.mdpi.com/1420-3049/26/24/7534 (2021).

28. Jeschke, G. *et al.* DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data. *Applied Magnetic Resonance* **30,** 473–498. http://link.springer.com/10.1007/BF03166213 (2006).

29. Hiscock, H. G. *et al.* The quantum needle of the avian magnetic compass. *Proceedings of the National Academy of Sciences* **113,** 4634–4639. https://pnas.org/doi/full/10.1073/pnas.1600341113 (Apr. 2016).

30. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nature Methods* **19,** 679–682. https://www.nature.com/articles/s41592-022-01488-1 (2022).

31. Prinz, W. A., Åslund, F., Holmgren, A. & Beckwith, J. The Role of the Thioredoxin and Glutaredoxin Pathways in Reducing Protein Disulfide Bonds in the Escherichia coliCytoplasm. *Journal of Biological Chemistry* **272,** 15661–15667. https://linkinghub.elsevier.com/retrieve/pii/S0021925818393797 (1997).

32. Jakob, U., Eser, M. & Bardwell, J. C. Redox Switch of Hsp33 Has a Novel Zinc-binding Motif. *Journal of Biological Chemistry* **275,** 38302–38310. http://dx.doi.org/10.1074/jbc.M005957200%20https://linkinghub.elsevier.com/retrieve/pii/S0021925819559580 (2000).

33. Choi, H.-J. *et al.* Structural Basis of the Redox Switch in the OxyR Transcription Factor. *Cell* **105,** 103–113. https://linkinghub.elsevier.com/retrieve/pii/S0092867401003002 (2001).

34. Palayam, M. *et al.* Structural insights into photoactivation of plant Cryptochrome-2. *Communications Biology* **4,** 28. http://dx.doi.org/10.1038/s42003-020-01531-x%20https://www.nature.com/articles/s42003-020-01531-x (2021).

35. Shao, K. *et al.* The oligomeric structures of plant cryptochromes. *Nature Structural & Molecular Biology* **27,** 480–488. http://www.nature.com/articles/s41594-020-0420-x (2020).

36. Rosenfeldt, G., Viana, R. M., Mootz, H. D., von Arnim, A. G. & Batschauer, A. Chemically Induced and Light-Independent Cryptochrome Photoreceptor Activation. *Molecular Plant* **1,** 4–14. https://linkinghub.elsevier.com/retrieve/pii/S1674205214603221 (2008).

37. Sang, Y. *et al.* N-Terminal Domain–Mediated Homodimerization Is Required for Photoreceptor Activity of Arabidopsis CRYPTOCHROME 1. *The Plant Cell* **17,** 1569–1584. https://academic.oup.com/plcell/article/17/5/1569-1584/6114536 (2005).

38. Wang, Q. & Lin, C. Mechanisms of Cryptochrome-Mediated Photoresponses in Plants. *Annual Review of Plant Biology* **71,** 103–129. https://www.annualreviews.org/doi/10.1146/annurev-arplant-050718-100300 (2020).

39. Schmalen, I. *et al.* Interaction of Circadian Clock Proteins CRY1 and PER2 Is Modulated by Zinc Binding and Disulfide Bond Formation. *Cell* **157,** 1203–1215. http://dx.doi.org/10.1016/j.cell.2014.03.057%20https://linkinghub.elsevier.com/retrieve/pii/S0092867414005352 (2014).

40. Kondrat, D. L. F., Struwe, W. B. & Benesch, J. L. P. *Structural Proteomics* Second Edi (ed Owens, R. J.) 349–371. https://link.springer.com/10.1007/978-1-4939-2230-7 (Springer New York, New York, NY, 2015).

41. Quantitative mass imaging of single biological macromolecules. *Science* **360,** 423–427. https://www.science.org/doi/10.1126/science.aar5839 (2018).

42. Lazar, I. J. & Lazar, I. S. *GelAnalyzer 19.1.* www.gelanalyzer.com.

43. Iacobucci, C. *et al.* A cross-linking/mass spectrometry workflow based on MS-cleavable cross-linkers and the MeroX software for studying protein structures and protein–protein interactions. *Nature Protocols* **13,** 2864–2889 (2018).

44. Shevchenko, A., Tomas, H., Havli, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols* **1,** 2856–2860. http://www.nature.com/articles/nprot.2006.468 (2006).

45. Pannier, M., Veit, S., Godt, A., Jeschke, G. & Spiess, H. W. Dead-Time Free Measurement of Dipole – Dipole Interactions between Electron Spins. *Journal of Magnetic Resonance* **213,** 331–340 (2000).

I've been feeling wobbly, lately
I've been feeling wobbly, so wobbly
Feeling like they lost me, can't find me
And they're always one step behind me

— from Ezra Furman's *Wobbly*

6

Complementing experimental structural data through molecular dynamics simulations

Contents

6.1 Introduction 213
6.2 Supporting HDX-MS data on *Dm*Cry 215
6.2.1 Results and Discussion 215
6.2.2 Conclusion 220
6.3 Methods 221
References 222

6.1 Introduction

The navigational sense of birds is of high interest in the field of magnetoreception, as discussed in other chapters of this thesis. However, since birds are not a good model organism as their genetic modification to elucidate the role of the putative magnetoreceptor remains impossible [1], attention is often turned to *Drosophila melanogaster*, the fruit fly, a versatile model organism of which many mutants exist. Additionally, the crystal structure of *Drosophila melanogaster* cryptochrome

(*Dm*Cry) [2] was solved before any avian cryptochrome structure was solved and consequently quite a few studies covering *Dm*Cry's experimental [3–7] and theoretical [8–10] examination have been published to date. Despite a recent study questioning the fly's ability to perceive the Earth's magnetic field [11], its cryptochromes still remain of interest as an important model protein and step towards elucidating the function of other species' cryptochromes. The radical pair mechanism discussed in detail in other chapters applies in a very similar manner to *Dm*Cry where the FAD cofactor is excited and provided with an electron from a nearby Trp residue followed by several steps of electron transfers along a Trp tetrad to yield a radical pair separated in space [12].

Even though the photochemistry of *Dm*Cry is well understood, the question of how the protein might transmit a photochemical signal and which conformational changes might occur along the way is not yet answered. Several studies indicate that a part of the protein, the C-terminal tail (CTT, residues 525-539 [2]), is of importance for signalling. A C-terminal alpha-helix docks in the region of the protein that corresponds to the DNA binding groove in photolyases [2] and becomes more exposed upon blue light irradiation [10, 13, 14]. Furthermore, MD simulations performed by Ganguly *et al.* revealed that the CTT of *Dm*Cry might undergo conformational changes as a results of FAD photoreduction in a manner that is dependent on the hydrogen bonding pattern of histidine 378, a residue separating the cofactor from the CTT [15]. Another recent study found that only the CTT without the FAD-binding domain of *Dm*Cry is sufficient for signal transmission of magnetic-field responses in cells. The authors argue that Cry might fulfil a role as a magnetotransducer rather than a magnetoreceptor [16]. However, no data covering dynamic changes of the whole magnetotransducer or magnetoreceptor protein structure upon light-activation has been published so far and HDX-MS is a technique ideally suited to probe for exactly this kind of information.

The previous DPhil students Shane A. Chandler and Katherine Buckton at the University of Oxford have acquired such data using different illumination conditions and their data will be presented in this chapter. However, their data set exhibited

a few gaps in the sequence coverage, including the areas covering the phosphate-binding loop (residues 249–263 [17]) and a part of the protrusion motif (residues 288–306 [17]). Therefore, MD simulations were implemented to fill in the gaps and provide a more detailed analysis of potential conformational changes of *Dm*Cry post light-activation at short (sub-microsecond) timescales.

This chapter will briefly present the *Dm*Cry experimental results obtained by Chandler and Buckton, but mainly focus on the MD results obtained for the same protein. All of the results combined are currently being drafted into a manuscript for publication. However, in comparison to previous chapters this chapter is rather brief due to the general focus of this work on mass spectrometry.

6.2 Supporting HDX-MS data on *Dm*Cry

MD simulations on *Dm*Cry were performed to find out more about changes to the protein structure on short (sub-microsecond) timescales. Given that under conditions of constant illumination, which were chosen for the HDX-MS experiments, the Trp radicals are expected to be rather short-lived (in the microsecond regime) [18], the only radical chosen to be present within the simulations is $FAD^{\bullet-}$. The fourth member of the Trp tetrad might get deprotonated in its radical state to give a more stable, neutral radical, but this also happens on a microsecond timescale [19] with radical lifetimes greater than 100 µs [18]. The radical pair is expected to recombine within a time frame of several milliseconds [20]. Therefore the Trp radicals are expected to be too short-lived to have notable influences on changes in deuterium uptake. $FAD^{\bullet-}$ does not protonate as is the case for proteins discussed in previous chapters as there is no proton-donor nearby the N5 atom of the middle isoalloxazine ring, the site that usually gets protonated in other proteins, neither does $FAD^{\bullet-}$ get further reduced [19].

6.2.1 Results and Discussion

Therefore, three light state (LS) MD simulations were set up to contain only the $FAD^{\bullet-}$ radical. For the three dark state (DS) simulations, both the cofactor and all

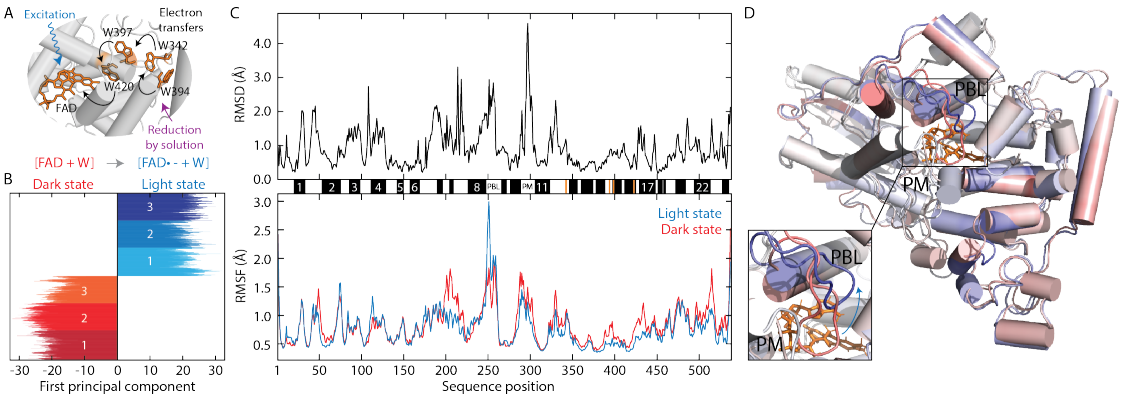

Figure 6.1: Conformational changes observed in molecular dynamics simulations. (**A**) Electron transfer pathway along the tryptophan tetrad following initial FAD excitation. Figure courtesy of Justin L. P. Benesch. (**B**) First principal component of three dark state simulations in different shades of red (prior to any electron transfer) and three light state simulations in different shades of blue (assuming a long-lived state with only the FAD anion radical being present). (**C**) Upper plot: RMSD differences between the most representative light and dark state structures. These structures were aligned in VMD to a curated subset of residues leaving out areas of interest (aligned to the backbone of residues 6-109, 131-247, and 305-484) prior to RMSD calculation. Lower plot: Average RMSF of three repeats of the dark and the light state. Secondary structure elements are depicted in between the two plots. (**D**) Overlay of the aligned most representative dark state and light state structures coloured with their residue's average RMSF values (each averaged over all three repeats). The colours for the dark state range from white to red and from white to blue for the light state. The inset shows a zoom into the region showing large RMSF values featuring the phosphate-binding loop (PBL, residues 249-263) and a part of the protrusion motif (PM, residues 288-306) as indicated. The PM extends further as indicated in the larger display.

Trp residues were set to be in their neutral ground states. All repeats of both states were simulated for 300 ns following a long, 400 ns equilibration period (initially, a 1 µs simulation was run for both states which did not yield significantly more information than a shorter simulation, so after a cutoff of the first 400 ns two more repeats of 300 ns were performed and used for the final analysis). The root-mean-square deviations (RMSD) were calculated, the most representative frames for both LS and DS simulations were determined and a principal component analysis (PCA) was performed. The results of this analysis are displayed in Figure 6.1. Panel A of the figure depicts schematically the electron transfer pathway during the radical formation process and panel B shows the results of the PCA: already the

first principal component separates the DS and LS repeats from each other which indicates that the two states have sufficiently different structural arrangements. The RMSD of the most representative structures (chosen to have the smallest eigenvalue deviations from the average eigenvalues of the first two principal components, see section 6.3 for further details) was calculated in order to find the differences between LS and DS simulations. As the upper part of panel C of the same figure shows, these differences are especially prominent for the protrusion motif (PM) and the phosphate-binding loop (PBL). The area between α7 and α8 also shows similarly large RMSD values for some residues, but not for the whole area. The lower plot of the same panel shows the root-mean-square fluctuation (RMSF) values of the most representative structures which are also projected onto a three-dimensional structure of the protein in panel D using the same colour coding. RMSF is a measure of the variation of amino acid positions with respect to their average position and therefore indicates the flexibility of different protein areas. The results shown here are very similar for the LS and the DS. One difference is that the PBL experiences larger fluctuations in the LS than in the DS. Other than that the dark state has slighter larger RMSF values across the sequences starting shortly after α6.

As Figure 6.2A displays, α22 shows much greater increase of uptake over an hour, from the shortest time point of 15 s to the longest of 1 hr, than any other part of the protein, even larger than what was observed for α23, the most C-terminal helix, in the DS of the protein. The LS shows some notable differences in uptake when comparing the results to the DS: α22 and α23 show a marked increase already after the shortest labelling time point of 15 s with α23 consistently showing a larger uptake than the DS at any given time point. A few other regions also show increased uptake after longer timescales covering α4, α16, and the loop between α14 and α15. The changes observed in these areas are not mirrored by the movements observed within the MD simulations. However, since the latter happen on much shorter timescales (sub-microseconds), it is conceivable that the uptake changes observed in the HDX-MS data would only be mirrored in MD simulations of much longer timescales. Panel C of Figure 6.2 has the change in uptake at different labelling

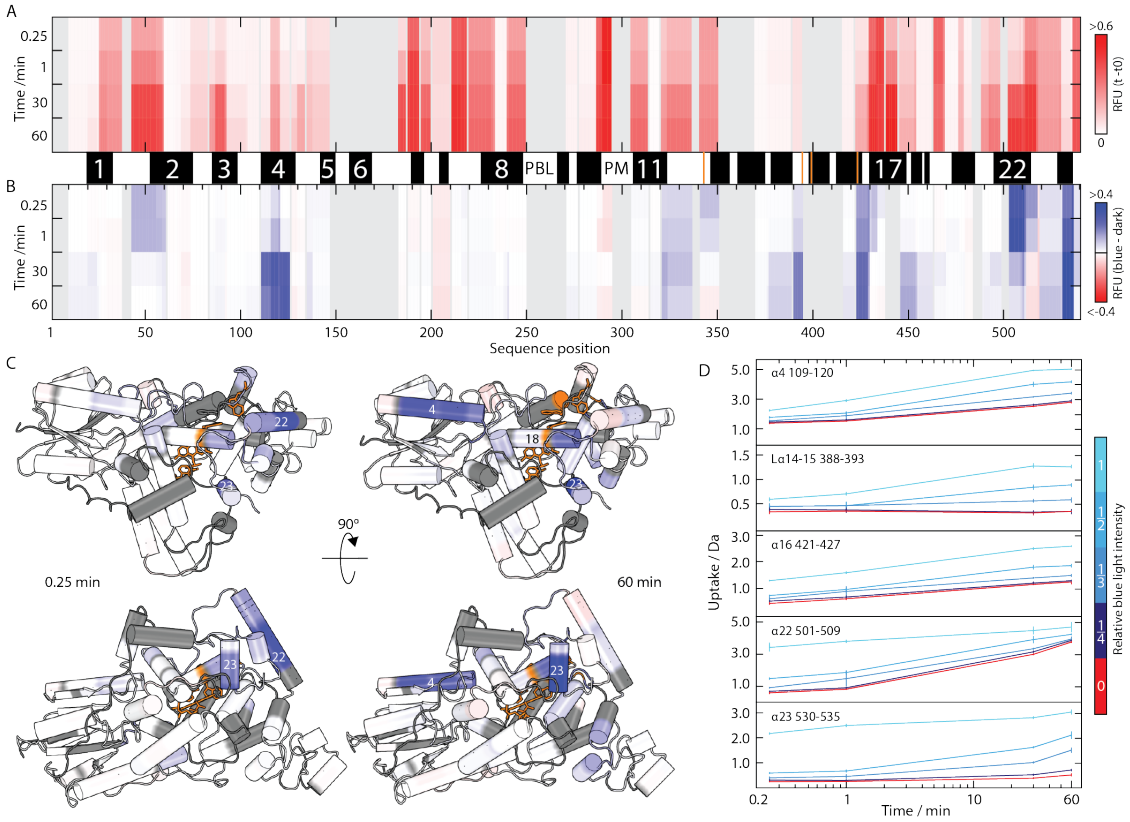

Figure 6.2: Conformational dynamics of *Dm*Cry WT observed by HDX-MS. (**A**) The degree of deuterium incorporation expressed as an uptake heat-map for *Dm*Cry under dark conditions. HDX-MS was performed on *Dm*Cry after exposure to a deuterated buffer solution for 0.25, 1, 30 and 60 min incubation periods. Estimates of the uptake at each residue are shown having been normalised to the theoretical maximum and correspond to values between 0 % (white) and 63 % (red) as indicated in the colour bar. A sequence coverage of 80.5 % was observed with regions of no coverage shown in grey. All α-helices are depicted above the sequence. (**B**) The change in deuterium uptake for WT *Dm*Cry upon the continuous exposure to blue light displayed as a differential heat-map. A positive change (blue) corresponds to greater uptake for the blue light environment. Regions which undergo no change in deuterium incorporation are shown in white and some residues undergo minor protection and are shown in red. (**C**) Uptake difference after 15 s of labelling mapped onto the *Dm*Cry crystal structure (PDB: 4GU5). (**D**) Representative mass spectra are shown for 2 peptides, 501-509 and 530-535, in the CTT region, α23 and α22 helices, which both show a large blue-light-induced increase in deuterium uptake. Dark spectra are shown in red with those exposed to blue light coloured in blue over the time course 0.25 - 60 min. In all panels, white regions indicate no coverage and the FAD/Trp-tetrad residues are shown in orange. Reduced blue-light intensity uptake curves are also shown as darker shades of blue. Figure and figure description courtesy of Shane A. Chandler and Justin L. P. Benesch. The original figure was adapted for clarity.

time points mapped onto the atomic coordinates of the protein DS. Regions of increased uptake are located near the FAD cofactor which roughly corresponds to the general area of increased flexibility found in the simulations (see Figure 6.1C, lower panel). Panel D of Figure 6.2 serves the purpose of validating that the light-induced changes occur due to the influence of the light by showing the increase in uptake for several peptides as a function of different light intensities.

Interestingly, the areas of *Dm*Cry that experience the most prominent changes between the different light conditions, are the ones that also stand out for the same reason for the bird cryptochromes investigated in chapter 4. The PBL and PM experience protection for the bird cryptochromes in the LS in comparison to the DS and are found to undergo changes in flexibility as well as displacement for *Dm*Cry, indicating that these areas might be important for signal transduction in all of the investigated cryptochromes. The absence of HDX-MS data within these regions for *Dm*Cry could stem from the increased flexibility in this area which might impact on the quality of the data acquired for peptides covering this region, ultimately rendering it too low for any valuable interpretation. An interesting difference between the HDX-MS data results for the fruit fly and the bird cryptochromes is that the C-terminal region seems to behave differently: in *Dm*Cry α22 experiences an increased deuterium uptake for the LS whereas for the bird cryptochromes this region experiences the opposite upon light-exposure. Another difference in the C-terminal region is that for the bird cryptochromes the very end of the C-terminal tail experiences very high deuterium uptake which is not the case to the same extent for *Dm*Cry. These differences in response to light-exposure might highlight how these proteins have evolved to fulfil different functions within the different species.

For *Dm*Cry, several binding motifs in the C-terminal region have been found, among which are a TRAF2 ligand motif and two class III PDZ-binding domains, the combination of which is known to be present in proteins which participate in multiple different signalling and interaction processes [21, 22]. One of the putative PDZ motifs (528-531) was recently found to be involved in a magnetic field effect observed in fruit flies as its disruption via mutation of a key amino acid prevented

a response to a magnetic field. However, light sensitivity was retained, even in complete absence of the PDZ motif [16]. It would be conceivable that therefore a part of the C-terminus could be important for magnetic field-induced signalling whereas the PHR region of the protein might be more important for light-dependent signalling. The here presented work suggests that the regions of importance for this could involve the PM and and the PBL as these are two regions of the PHR which display changes in the LS.

The experiments by Chandler and Buckton further included HDX-MS results performed on the W342F and the W394F *Dm*Cry mutants where the electron transfer chain was blocked with a redox-inactive phenylalanine in the positions of the second to last and last member of the Trp tetrad, respectively. However, analysing the results would go beyond the scope of this thesis and will be discussed in detail in a manuscript in preparation covering these and the here explored results.

6.2.2 Conclusion

The aim of this chapter was to explore how MD simulations could complement experimental HDX-MS results of the photoinduced changes occurring in *Dm*Cry, a model protein for the putative magnetoreceptor in migratory birds. The here presented work does not cover any magnetic-field induced changes, but photo-induced changes on timescales where it is assumed that the only radical present is $FAD^{\bullet-}$.

The light-induced changes found towards the C-terminal region in the HDX-MS data mirror other findings in the literature that the CTT becomes more accessible upon photoexcitation. The reason for these not being observed in the MD simulations is likely to be due to the different timescales of the two approaches. However, the simulations were able to reveal changes occurring in the regions that were not covered by the HDX-MS data, specifically for the PBL and the PM. The finding that these areas show differences in flexibility and also in displacement between the LS and the DS are interesting, also in light of the findings presented for avian Cry4a proteins presented in chapter 4. For the *Dm*Cry LS the flexibility increases for the PBL and decreases for the PM at sub-microsecond timescales with

the largest overall change in displacement between the LS and the DS occurring for the PM. Interestingly, this is also the region where a lot of overlapping peptides in avian Cry4a proteins show significant changes in surface accessibility.

Overall, the MD data presented here show how simulations can supplement HDX-MS experiments, as in the work by Kattnig *et al.* in 2018 [17] which parallels the experimental findings presented in chapter 4, and provide deeper insights into light-induced structural changes in proteins.

6.3 Methods

A three-dimensional structure of *Dm*Cry1 was built based on the PDB entry 4GU5 and previously used elsewhere [23]. All simulations were performed using the NAMD molecular dynamics (MD) package [24] and the CHARMM36 force field [25–27].

The first 140 ns of the already published dark state of the protein (no radicals present) was extended. The 400th nanosecond of this simulation was used as a starting point for three individual repeats lasting about 300 ns. Small variations in the exact length of the simulations (some are up to 42 ns longer) occurred due to slightly different lengths of the initial simulations that the present simulations are based on.

For the light state, a first simulation based on the already ionised and neutralised dark state version of the protein used elsewhere [23] was performed after switching the cofactor FAD to its radical anion version ($FAD^{\bullet-}$ as parameterised previously [28]) and neutralising the model. A patch protonating aspartic acid residue 31 which was used in the dark state simulation published elsewhere [23] was not applied in this case. The model was equilibrated for 10 ns with a constrained backbone, then dynamically equilibrated for 20 ns without constraints followed by 400 ns of a production simulation. This was subsequently taken as the starting point for three repeats lasting 300 ns each.

All simulations were performed with an integration time step of 2 fs at 310 K using the Langevin thermostat with a damping coefficient of 5.0 ps^{-1} applied to all atoms except hydrogens to keep the temperature stable. The simulations were performed using the NPT ensemble where the pressure and temperature were held constant.

The Langevin barostat [29] was used to keep the pressure constant at 1 atm. Periodic boundary conditions were implemented, the Particle Mesh Ewald summation method [30] was used to evaluate Coulomb forces, and the ShakeH algorithm was used to keep the bonds to hydrogens at a fixed length. Using a cut-off distance of 12 Å and a switching distance of 10 Å the van der Waals energy was calculated.

The MD results were analysed using VMD 1.9.3 [31] and custom-written Python scripts. All RMSD values were calculated based on a backbone alignment excluding the areas in which large deuteration changes in the HDX-MS data upon photoexcitation were observed. Therefore, the backbone was aligned to residues 6-109, 131-247, and 305-484 of the first frame of the simulation. Every tenth frame was subjected to the analysis of both RMSD and RMSF values.

PCA was performed using the scikit-klearn tool within Python based on a code written by Guodong Cao, a DPhil student in the Benesch group at the University of Oxford. Briefly, one frame for each ns of the production run was sampled for the analysis. The analysis was based on the extraction of all Cα coordinates. Their dimension was reduced by extracting the first two principal components, PC1 and PC2, and their corresponding eigenvalues. The eigenvalues of all structures, corresponding to PC1 and PC2 were averaged independently. To determine the most representative structure of the series, the structure which had the smallest eigenvalue deviation from the average eigenvalues calculated based on the weight of PC1 and PC2 was determined. This was again done for the dark and the light states separately.

The experimental methods used to obtain the data displayed above are described elsewhere (manuscript in preparation).

References

1. Kyriacou, C. P. & Rosato, E. Genetic analysis of cryptochrome in insect magnetosensitivity. *Frontiers in Physiology* **13,** 1–14. https://www.frontiersin.org/articles/10.3389/fphys.2022.928416/full (2022).

2. Zoltowski, B. D. *et al.* Structure of full-length Drosophila cryptochrome. *Nature* **480,** 396–399. https://www.nature.com/articles/nature10618 (2011).

3. Foley, L. E., Gegear, R. J. & Reppert, S. M. Human cryptochrome exhibits light-dependent magnetosensitivity. *Nature Communications* **2,** 356. https://www.nature.com/articles/ncomms1364 (2011).

4. Gegear, R. J., Casselman, A., Waddell, S. & Reppert, S. M. Cryptochrome mediates light-dependent magnetosensitivity in Drosophila. *Nature* **454,** 1014–1018 (2008).

5. Gegear, R. J., Foley, L. E., Casselman, A. & Reppert, S. M. Animal cryptochromes mediate magnetoreception by an unconventional photochemical mechanism. *Nature* **463,** 804–807. https://www.nature.com/articles/nature08719 (2010).

6. Fedele, G., Green, E. W., Rosato, E. & Kyriacou, C. P. An electromagnetic field disrupts negative geotaxis in Drosophila via a CRY-dependent pathway. *Nature Communications* **5,** 4391. http://www.nature.com/articles/ncomms5391 (2014).

7. Fedele, G. *et al.* Genetic Analysis of Circadian Responses to Low Frequency Electromagnetic Fields in Drosophila melanogaster. *PLoS Genetics* **10** (ed Taghert, P. H.) e1004804. https://dx.plos.org/10.1371/journal.pgen.1004804 (2014).

8. Solov'yov, I. A., Domratcheva, T. & Schulten, K. Separation of photo-induced radical pair in cryptochrome to a functionally critical distance. *Scientific Reports* **4** (2014).

9. Hong, G. & Pachter, R. Photoactivation of cryptochromes from drosophila melanogaster and sylvia borin: Insight into the chemical compass mechanism by computational investigation. *Journal of Physical Chemistry B* **119,** 3883–3892 (2015).

10. Berntsson, O. *et al.* Photoactivation of Drosophila melanogaster cryptochrome through sequential conformational transitions. *Science Advances* **5,** eaaw1531. https://advances.sciencemag.org/lookup/doi/10.1126/sciadv.aaw1531%20https://www.science.org/doi/10.1126/sciadv.aaw1531 (2019).

11. Bassetto, M. *et al.* No evidence for magnetic field effects on the behaviour of Drosophila. *Nature* **620,** 595–599. https://www.nature.com/articles/s41586-023-06397-7 (2023).

12. Bradlaugh, A., Munro, A. L., Jones, A. R. & Baines, R. A. Exploiting the Fruitfly, Drosophila melanogaster, to Identify the Molecular Basis of Cryptochrome-Dependent Magnetosensitivity. *Quantum Reports* **3,** 127–136. https://www.mdpi.com/2624-960X/3/1/7 (2021).

13. Vaidya, A. T. *et al.* Flavin reduction activates Drosophila cryptochrome. *Proceedings of the National Academy of Sciences* **110,** 20455–20460. https://pnas.org/doi/full/10.1073/pnas.1313336110 (2013).

14. Chandrasekaran, S. *et al.* Tuning flavin environment to detect and control light-induced conformational switching in Drosophila cryptochrome. *Communications Biology* **4,** 249. http://dx.doi.org/10.1038/s42003-021-01766-2%20http://www.nature.com/articles/s42003-021-01766-2%20https://www.nature.com/articles/s42003-021-01766-2 (2021).

15. Ganguly, A. *et al.* Changes in active site histidine hydrogen bonding trigger cryptochrome activation. *Proceedings of the National Academy of Sciences* **113,** 10073–10078. https://pnas.org/doi/full/10.1073/pnas.1606610113 (2016).

16. Bradlaugh, A. A. *et al.* Essential elements of radical pair magnetosensitivity in Drosophila. *Nature* **615,** 111–116. https://www.biorxiv.org/content/10.1101/2021.10.29.466426v1%0Ahttps://www.biorxiv.org/content/10.1101/2021.10.29.466426v1.abstract%20https://www.nature.com/articles/s41586-023-05735-z (2023).

17. Kattnig, D. R., Nielsen, C. & Solov'yov, I. A. Molecular dynamics simulations disclose early stages of the photo-activation of cryptochrome 4. *New Journal of Physics* **20,** 083018. https://iopscience.iop.org/article/10.1088/1367-2630/aad70f (2018).

18. Sheppard, D. M. *et al.* Millitesla magnetic field effects on the photocycle of an animal cryptochrome. *Scientific Reports* **7,** 1–7 (2017).

19. Paulus, B. *et al.* Spectroscopic characterization of radicals and radical pairs in fruit fly cryptochrome – protonated and nonprotonated flavin radical-states. *The FEBS Journal* **282,** 3175–3189. https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.13299 (2015).

20. Einholz, C. *et al.* pH-dependence of signaling-state formation in Drosophila cryptochrome. *Archives of Biochemistry and Biophysics* **700,** 108787. https://doi.org/10.1016/j.abb.2021.108787%20https://linkinghub.elsevier.com/retrieve/pii/S0003986121000370 (2021).

21. Hemsley, M. J. *et al.* Linear motifs in the C-terminus of D. melanogaster cryptochrome. *Biochemical and Biophysical Research Communications* **355,** 531–537. https://linkinghub.elsevier.com/retrieve/pii/S0006291X07002690 (2007).

22. Mazzotta, G. M. *et al.* Calmodulin Enhances Cryptochrome Binding to INAD in Drosophila Photoreceptors. *Frontiers in Molecular Neuroscience* **11,** 1–15. https://www.frontiersin.org/article/10.3389/fnmol.2018.00280/full (2018).

23. Nielsen, C., Kattnig, D. R., Sjulstok, E., Hore, P. J. & Solov'yov, I. A. Ascorbic acid may not be involved in cryptochrome-based magnetoreception. *Journal of the Royal Society Interface* **14** (2017).

24. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **26,** 1781–1802 (2005).

25. MacKerell, A. D. *et al.* All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *The Journal of Physical Chemistry B* **102,** 3586–3616. https://pubs.acs.org/doi/10.1021/jp973084f (1998).

26. Mackerell, A. D., Feig, M. & Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry* **25,** 1400–1415. https://onlinelibrary.wiley.com/doi/10.1002/jcc.20065 (2004).

27. Best, R. B. *et al. Journal of Chemical Theory and Computation* **8,** 3257–3273. https://pubs.acs.org/doi/10.1021/ct300400x (2012).

28. Lüdemann, G., Solov'yov, I. A., Kubař, T. & Elstner, M. Solvent Driving Force Ensures Fast Formation of a Persistent and Well-Separated Radical Pair in Plant Cryptochrome. *Journal of the American Chemical Society* **137,** 1147–1156. https://pubs.acs.org/doi/10.1021/ja510550g (2015).

29. Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *The Journal of Chemical Physics* **103,** 4613–4621. http://aip.scitation.org/doi/10.1063/1.470648 (1995).

30. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N · log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **98,** 10089–10092. http://aip.scitation.org/doi/10.1063/1.464397 (1993).

31. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **14,** 33–38. arXiv: 1503.05249. https://linkinghub.elsevier.com/retrieve/pii/0263785596000185 (1996).

But, you know, you never trusted tomorrow
Oh, baby, that's just a problem of mine
Yeah, I've been thrown around by some bad ones, girl
Oh, and a good one was a new thing this time
Yeah, but you showed me that I couldn't be farther
From the truth of the problem at hand

— from Shakey Graves' *Tomorrow*

7

Conclusions and future work

The work described in this thesis aims to elucidate protein conformational changes involved in signal transduction using HDX-MS for avian cryptochromes and MD to complement previous HDX-MS data on *Dm*Cry. One main finding resulting from this investigation is that two areas of all proteins investigated seem to be of major importance: the phosphate-binding loop and the protrusion motif. For *Dm*Cry, these regions were not covered by the HDX-MS data but the MD simulations showed increased flexibility and movements in these areas. All avian cryptochromes, although many differences between species were found, show that these two areas become less solvent exposed upon blue-light irradiation, one possible interpretation being that the two regions might protect each other from the surrounding medium. Further interesting regions that experience consistently significant changes upon light-exposure are the already at short labelling time points rather solvent-exposed CTT and the region between residues 404-426. Again, one possible interpretation, given the experiment was performed on protein monomers, could be that the region covering 404-426 and the CTT protect each other. The PBL and the CTT were previously also found by Zoltowski *et al.* to experience light-induced protection in a limited proteolysis assay [1] which our results corroborate and expand on. With these findings this work lays the foundation for further explorations into

the question of how cryptochrome signal transduction might function *in vivo*. We have shown what happens to the protein and its separate regions, but the question as to which function these changes fulfil remains open. One suggestion, which is partially based on the work on *Dm*Cry by Bradlaugh *et al.*, [2] is that the changes observed within the PHR might mediate the light-response while the CTT might be more important and involved in a magnetic field-dependent response. However, as this work does not show convincing evidence for the latter, this is just a speculation and requires further work.

One possible explanation might be that conformational changes might happen as a consequence of photoexcitation and radical pair formation to enable the avian cryptochromes to form functional dimers. Although this thesis cannot answer questions of function, we have demonstrated that dimers are formed only in the presence of the cofactor FAD upon light-excitation and we have argued that at least a fraction of these dimers is covalently linked via disulphide bonds, with C317 and C412 being identified as involved in the dimerisation process for *Er*Cry4a. This work cannot exclude the presence of non-covalently bounds dimers. On the contrary, our collaborators have found several stable non-covalently formed dimers while performing their simulations. Further investigations are necessary to find out more about the extent to which covalently and non-covalently formed dimers exist and what functions they might fulfil. Interestingly, *At*Cry1 has been found to dimerise mediated via the N-terminal domain and a structural rearrangement of the dimer upon light-activation has been suggested [3]. It would be conceivable that the here investigated avian cryptochromes might function in a similar manner, although currently it seems more likely that light-induced conformational changes, potentially those within the PHR domain, happen before and promote dimerisation. In 2016 Wang *et al.* found that *At*Cry2 homodimerises in a light-dependent manner to become physiologically active which is in this case additionally regulated by BIC1, a protein found to bind to Cry2 and thereby inhibit its dimerisation processes [4]. Although different functions have been found for different cryptochromes, this suggests the possibility that other cryptochromes might also have evolved to become

physiologically active after light exposure and once they have dimerised. It further suggests that the possibility that dimerisation is regulated by downstream interaction partners should be taken into account for the design of future experiments. Recently, the tetrameric cryo-EM structures of two plant cryptochromes, a constitutively active *At*Cry2 mutant and a *Zea mays* Cry1 mutant, containing a mutation at the corresponding position, were solved [5]. Their dimerisation surfaces do not directly correspond to what the data presented here suggest for the avian cryptochromes. However, one of their interfaces within the PHR domain is situated close to the CTT which takes the interface to a similar area to where chemical crosslinks were found between two *Er*Cry4a monomer units in this work. However, to properly compare the avian with the plant cryptochrome dimerisation interface, further experiments are needed, e.g. the mixed approach, described in chapter 5, of using HDX-MS to elucidate the interface region and DEER-EPR to gain information on the orientation of the monomer units with respect to one another.

One initial aim of this work was also to investigate magnet-induced changes in solvent exposure using HDX-MS. Although some significant peptides displaying a different solvent-exposure with and without a magnet present were found, no general conclusions can be drawn from our data with respect to the responsiveness to an additional, external magnetic field. Magnet-induced changes were only expected to be present when the protein was also photoexcited. However, we also found a few areas of significant changes upon magnet-exposure in the dark which unfortunately renders our findings for the light state somewhat less reliable. Our significant peptides for these comparisons are therefore rather statistical artefacts and we were not able to find magnetic field effects under the experimental regime that we probed.

Despite not having been able to find definite magnet-induced changes, we were able to demonstrate in a detailed manner multiple changes occuring in the proteins upon light-excitation of which the consistently found dimerisation is a somewhat unexpected result opening up new avenues for research projects. Given that cryptochromes from other species form functional oligomeric units, the question arises whether the cryptochrome dimer might also fulfil a specific function,

potentially as part of a downstream signalling cascade. It would be conceivable that it might be more or less prone to protein-protein interactions in its dimeric as opposed to its monomeric state which might influence its interactions with its putative interaction partners that have recently been found [6, 7]. However, it should also be taken into consideration that higher-order oligomers such as trimers and tetramers are also being formed, although to a lesser extent than the dimer. Nevertheless, they might have a biological function where dimerisation might be the nucleation event to facilitate larger complexes or chains of cryptochromes which might even contribute to an alignment within the cell that is thought to be required to fulfil a function as a magnetoreceptor. Also of interest is the question of whether light-induced conformational changes might influence Cry4a's interaction with its interaction partners assuming it would do so in its monomeric state. This could be further investigated using HDX-MS which could yield information on an interaction surface.

Using DEER a first attempt at elucidating the interaction surface between two monomeric Cry4a units was made which could be expanded to different mutants and other protein species to further narrow down the interaction surface. Additionally, a successful HDX-MS experiment on cryptochrome dimers might lead to the information necessary to allow for an improved interpretation of the DEER data and to elucidate the dimer configuration.

Hopefully, the work presented here can provide new ideas for future experiments based on the findings of conformational changes and dimerisation as potential steps in the signal transduction cascade ultimately delivering perceived, potentially navigational information, to the brain of a living organism.

References

1. Zoltowski, B. D. *et al.* Chemical and structural analysis of a photoactive vertebrate cryptochrome from pigeon. *Proceedings of the National Academy of Sciences* **116,** 19449–19457. http://www.pnas.org/lookup/doi/10.1073/pnas.1907875116 (2019).

2. Bradlaugh, A. A. *et al.* Essential elements of radical pair magnetosensitivity in Drosophila. *Nature* **615,** 111–116. https://www.biorxiv.org/content/10.1101/2021.10.29.466426v1%0Ahttps://www.biorxiv.org/content/10.1101/2021.10.29.466426v1.abstract%20https://www.nature.com/articles/s41586-023-05735-z (2023).

3. Sang, Y. *et al.* N-Terminal Domain–Mediated Homodimerization Is Required for Photoreceptor Activity of Arabidopsis CRYPTOCHROME 1. *The Plant Cell* **17,** 1569–1584. https://academic.oup.com/plcell/article/17/5/1569-1584/6114536 (2005).

4. Wang, Q. *et al.* Photoactivation and inactivation of Arabidopsis cryptochrome 2. *Science* **354,** 343–347. https://www.sciencemag.org/lookup/doi/10.1126/science.aaf9030 (2016).

5. Shao, K. *et al.* The oligomeric structures of plant cryptochromes. *Nature Structural & Molecular Biology* **27,** 480–488. http://www.nature.com/articles/s41594-020-0420-x (2020).

6. Wu, H., Scholten, A., Einwich, A., Mouritsen, H. & Koch, K.-W. Protein-protein interaction of the putative magnetoreceptor cryptochrome 4 expressed in the avian retina. *Scientific Reports* **10,** 7364. http://www.nature.com/articles/s41598-020-64429-y%20https://www.nature.com/articles/s41598-020-64429-y (2020).

7. Görtemaker, K. *et al.* Direct Interaction of Avian Cryptochrome 4 with a Cone Specific G-Protein. *Cells* **11,** 2043. https://www.mdpi.com/2073-4409/11/13/2043 (2022).

Appendices

A

Appendix

Contents

A.1 Appendix for chapter 3 236
A.1.1 Peptide coverage for *Er*Cry4a WT and *Er*Cry4a W369F. 236
A.1.2 Uptake curves for *Er*Cry4a W369F peptides of interest. 237
A.1.3 Uptake curves on a logarithmic time scale for *Er*Cry4a W369F peptides of interest. 241
A.2 Appendix for chapter 4 245
A.2.1 Relative fractional uptake 245
A.2.2 Chiclet plots created in HD-eXplosion 249
A.2.3 ΔRFU plots created by Laila M. N. Shah 256
A.3 Appendix for chapter 5 271

A.1 Appendix for chapter 3

A.1.1 Peptide coverage for *Er*Cry4a WT and *Er*Cry4a W369F.

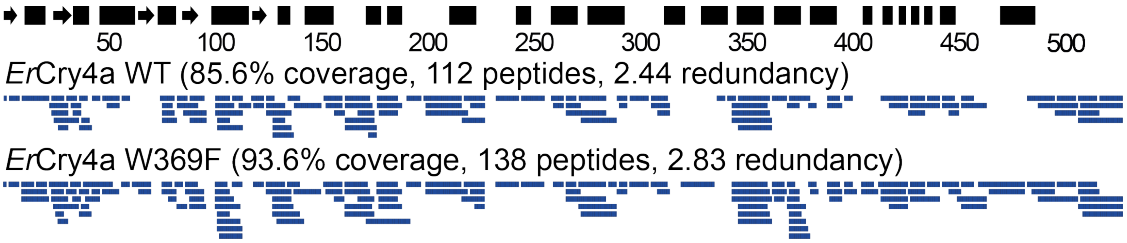

Figure A.1: Coverage maps for *Er*Cry4a WT and *Er*Cry4a W369F. Coverage maps after analysis in the DynamX 3.0 software (Waters). Beta sheets (arrows) and alpha-helices (rectangles) are indicated above the sequence.

A.1.2 Uptake curves for *Er*Cry4a W369F peptides of interest.

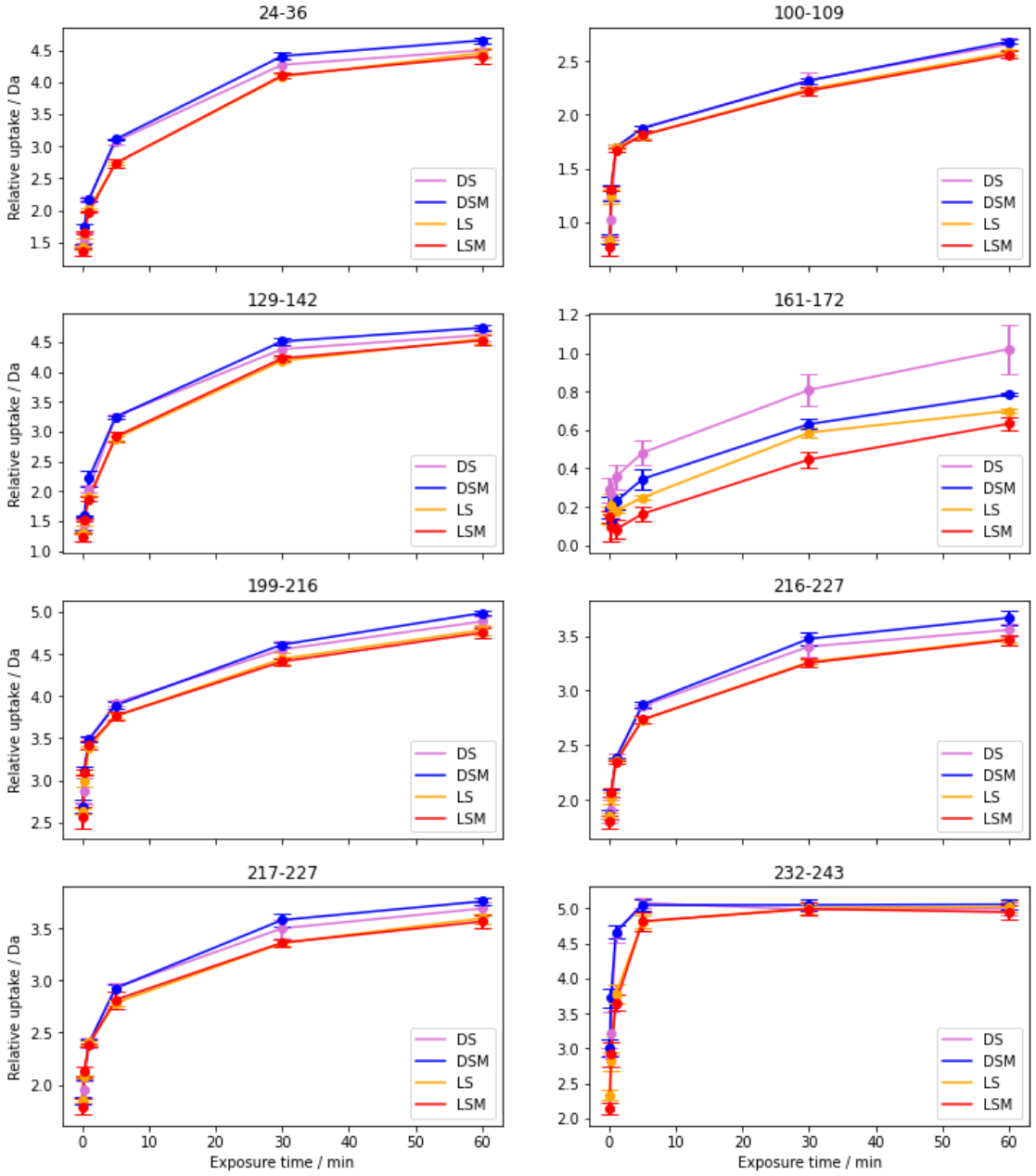

Figure A.2: Relative uptake of *Er*Cry4a W369F peptides of interest. Relative uptake of *Er*Cry4a W369F peptides that experience light-induced conformational changes over time. For more information see chapter 3. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

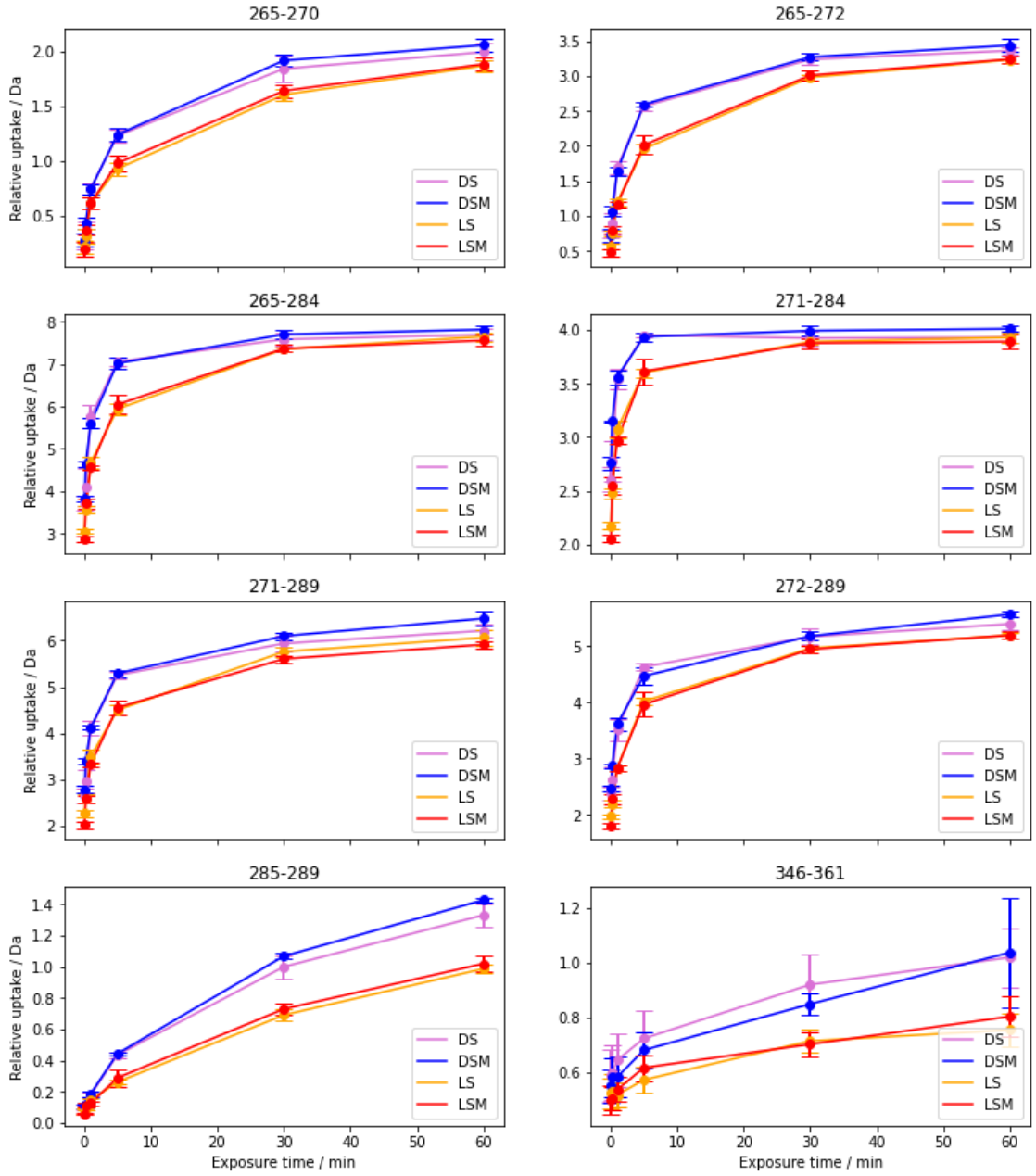

Figure A.3: Relative uptake of *Er*Cry4a W369F peptides of interest. Relative uptake of *Er*Cry4a W369F peptides that experience light-induced conformational changes over time. For more information see chapter 3. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

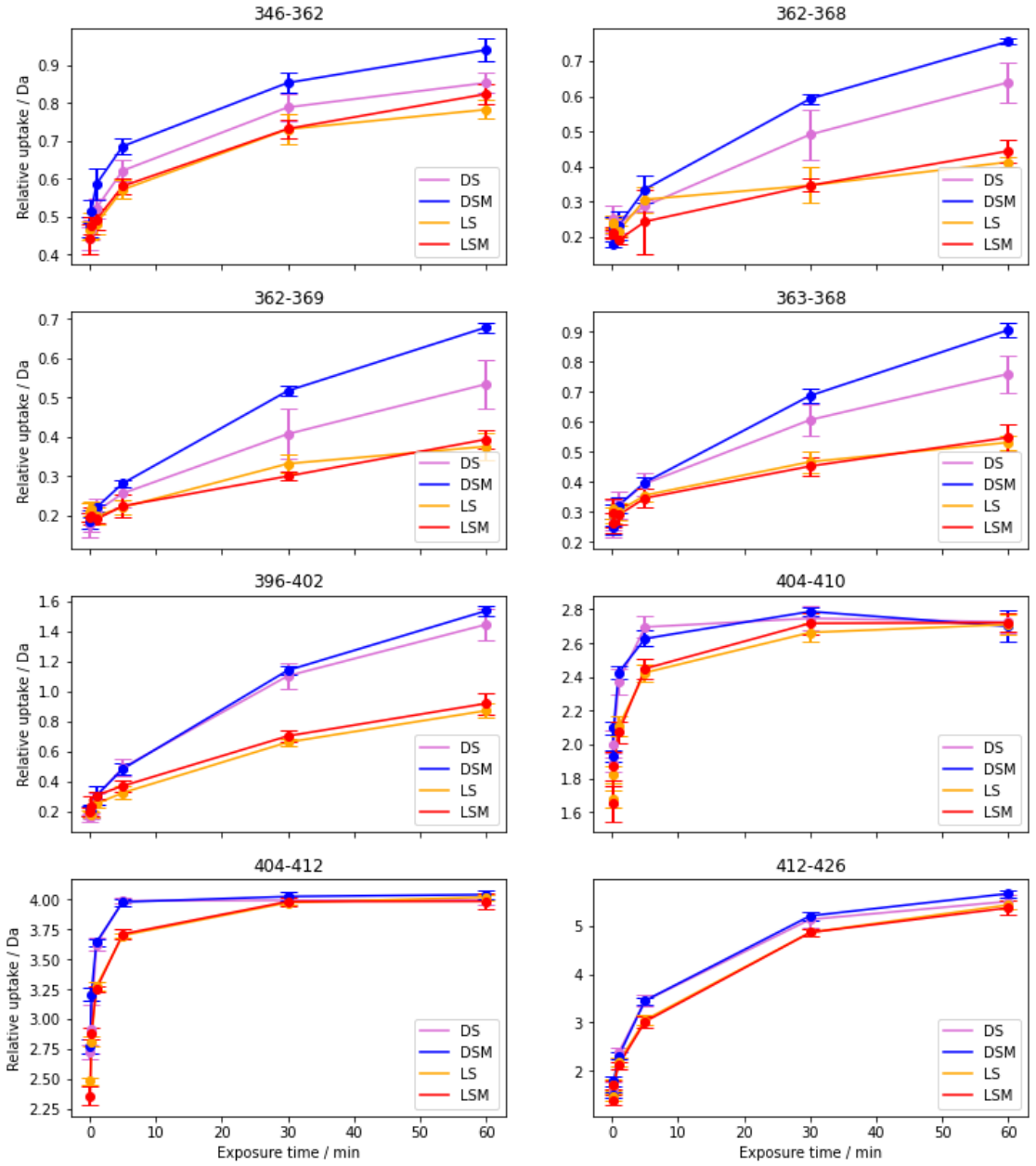

Figure A.4: Relative uptake of *Er*Cry4a W369F peptides of interest. Relative uptake of *Er*Cry4a W369F peptides that experience light-induced conformational changes over time. For more information see chapter 3. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

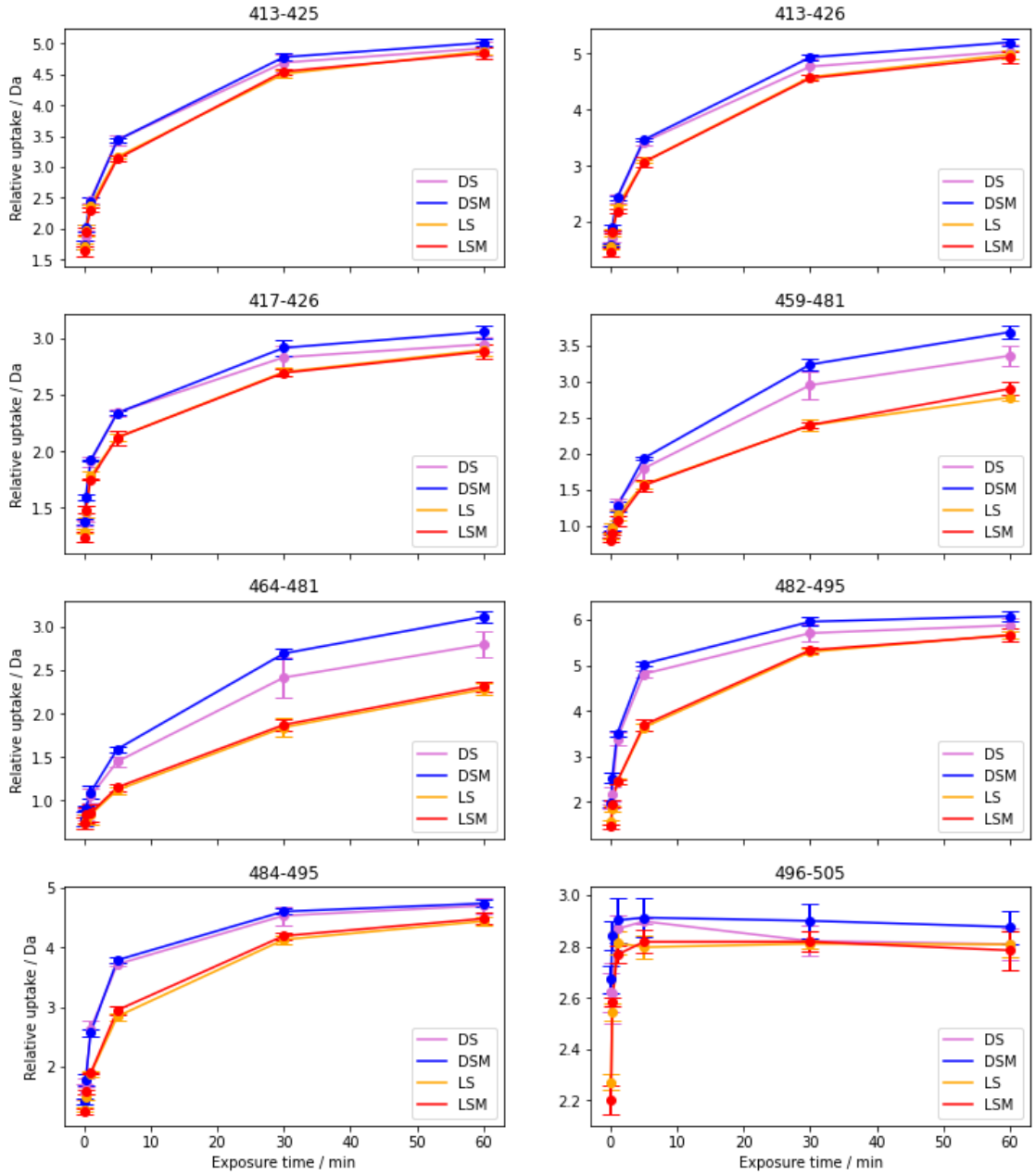

Figure A.5: Relative uptake of *Er*Cry4a W369F peptides of interest. Relative uptake of *Er*Cry4a W369F peptides that experience light-induced conformational changes over time. For more information see chapter 3. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

A.1.3 Uptake curves on a logarithmic time scale for *Er*Cry4a W369F peptides of interest.

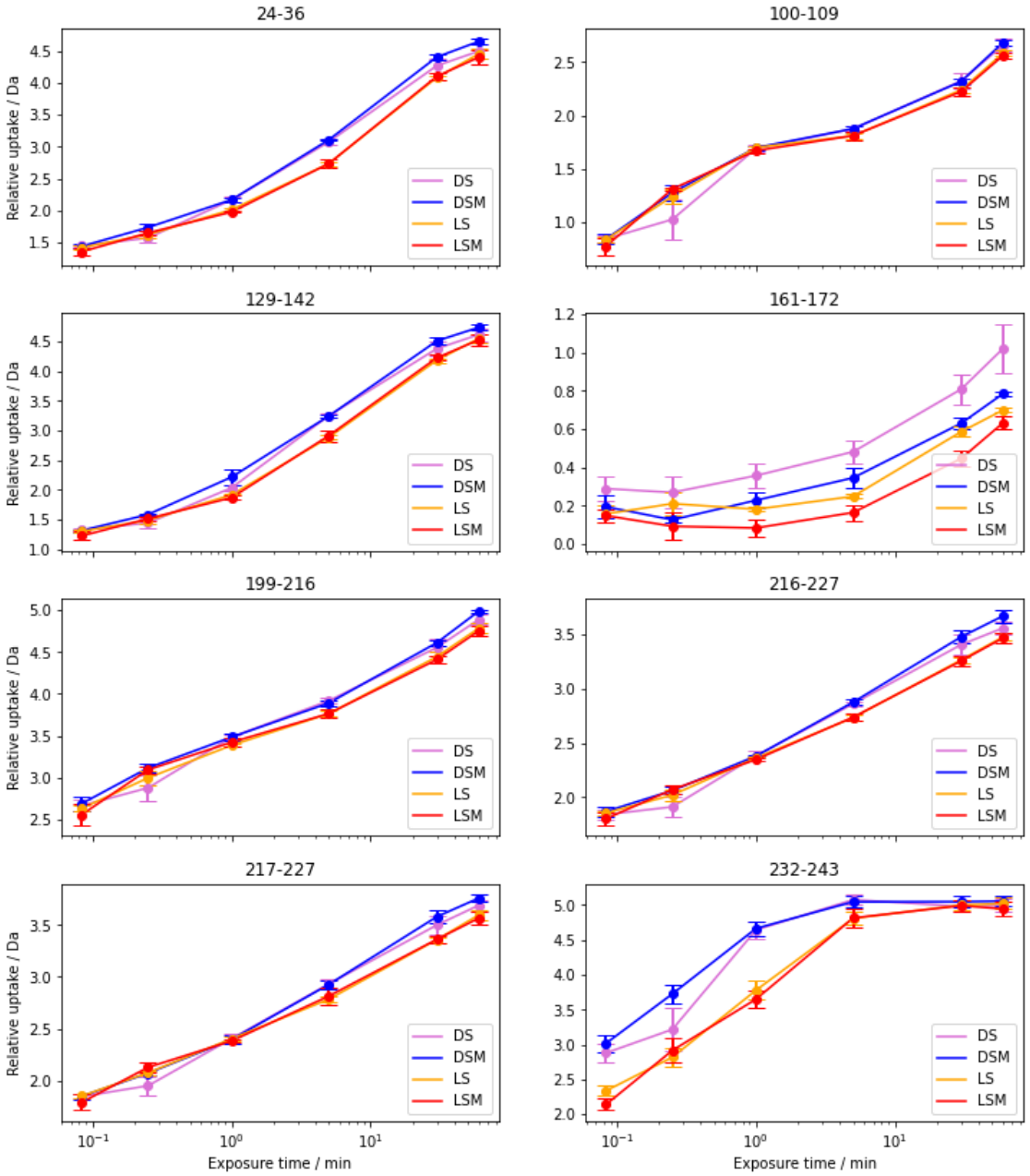

Figure A.6: Relative uptake of *Er*Cry4a W369F peptides of interest. Relative uptake of *Er*Cry4a W369F peptides that experience light-induced conformational changes over time displayed on a logarithmic time scale. For more information see chapter 3. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

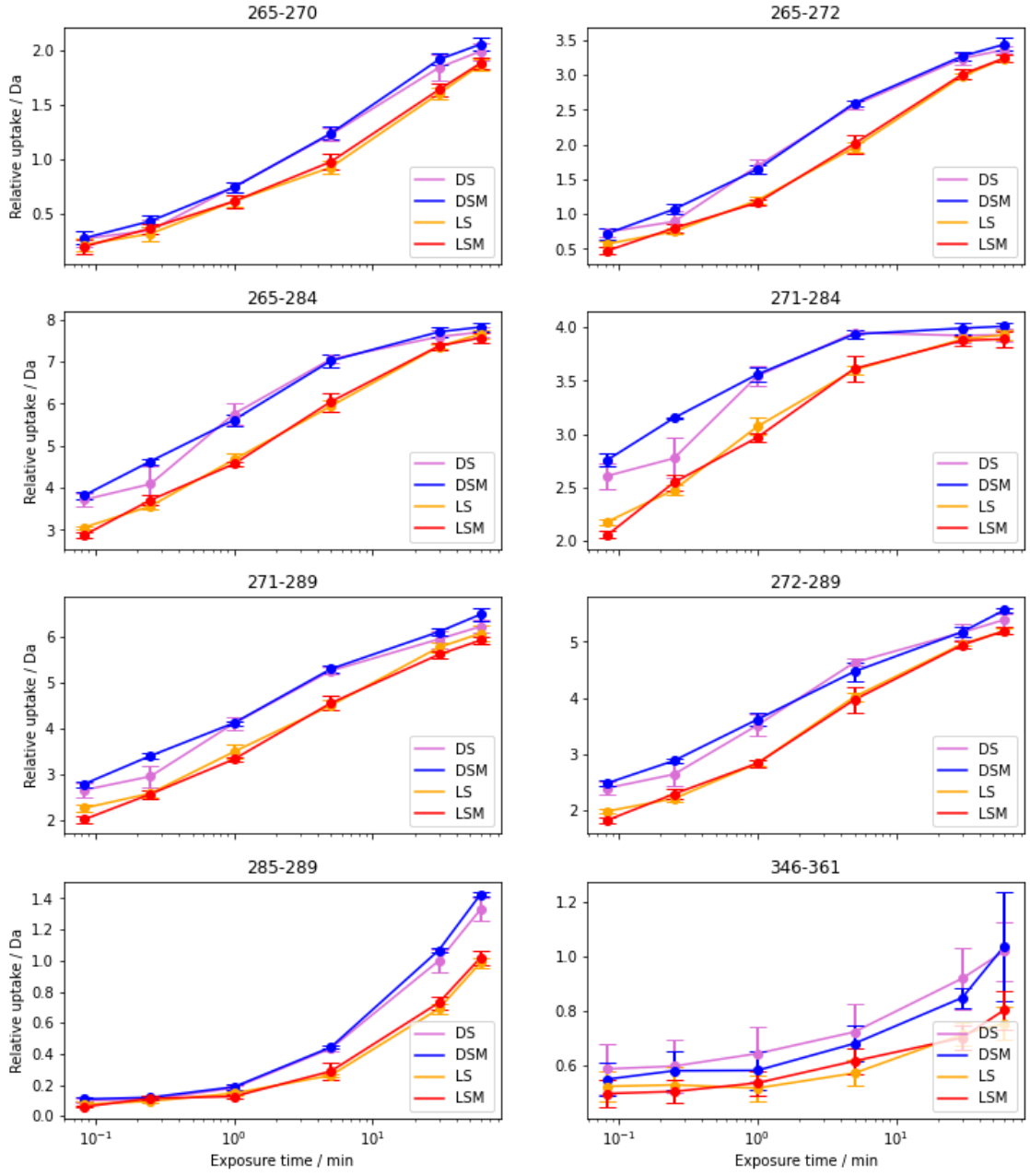

Figure A.7: Relative uptake of *Er*Cry4a W369F peptides of interest. Relative uptake of *Er*Cry4a W369F peptides that experience light-induced conformational changes over time displayed on a logarithmic time scale. For more information see chapter 3. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

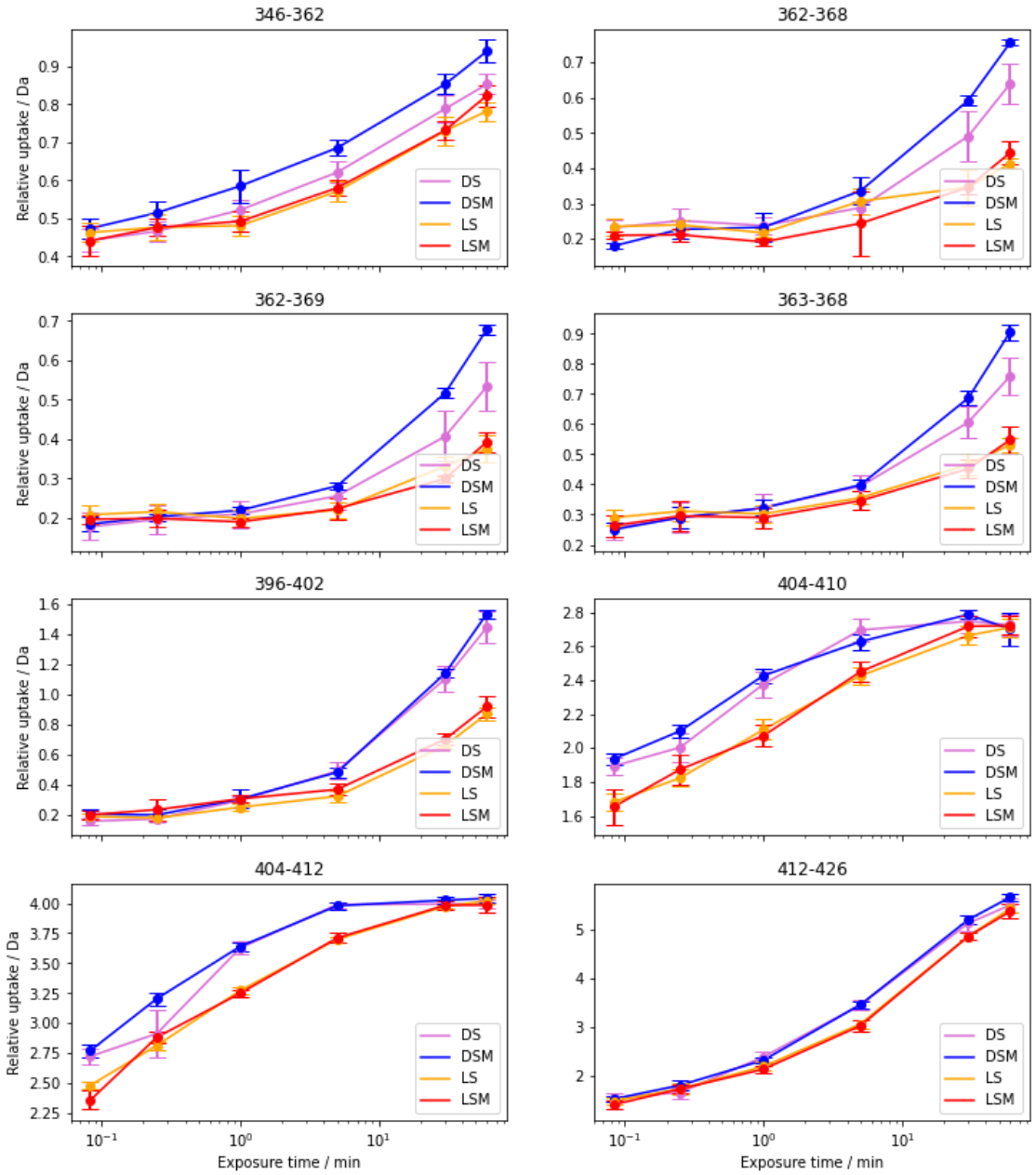

Figure A.8: Relative uptake of *Er*Cry4a W369F peptides of interest. Relative uptake of *Er*Cry4a W369F peptides that experience light-induced conformational changes over time displayed on a logarithmic time scale. For more information see chapter 3. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

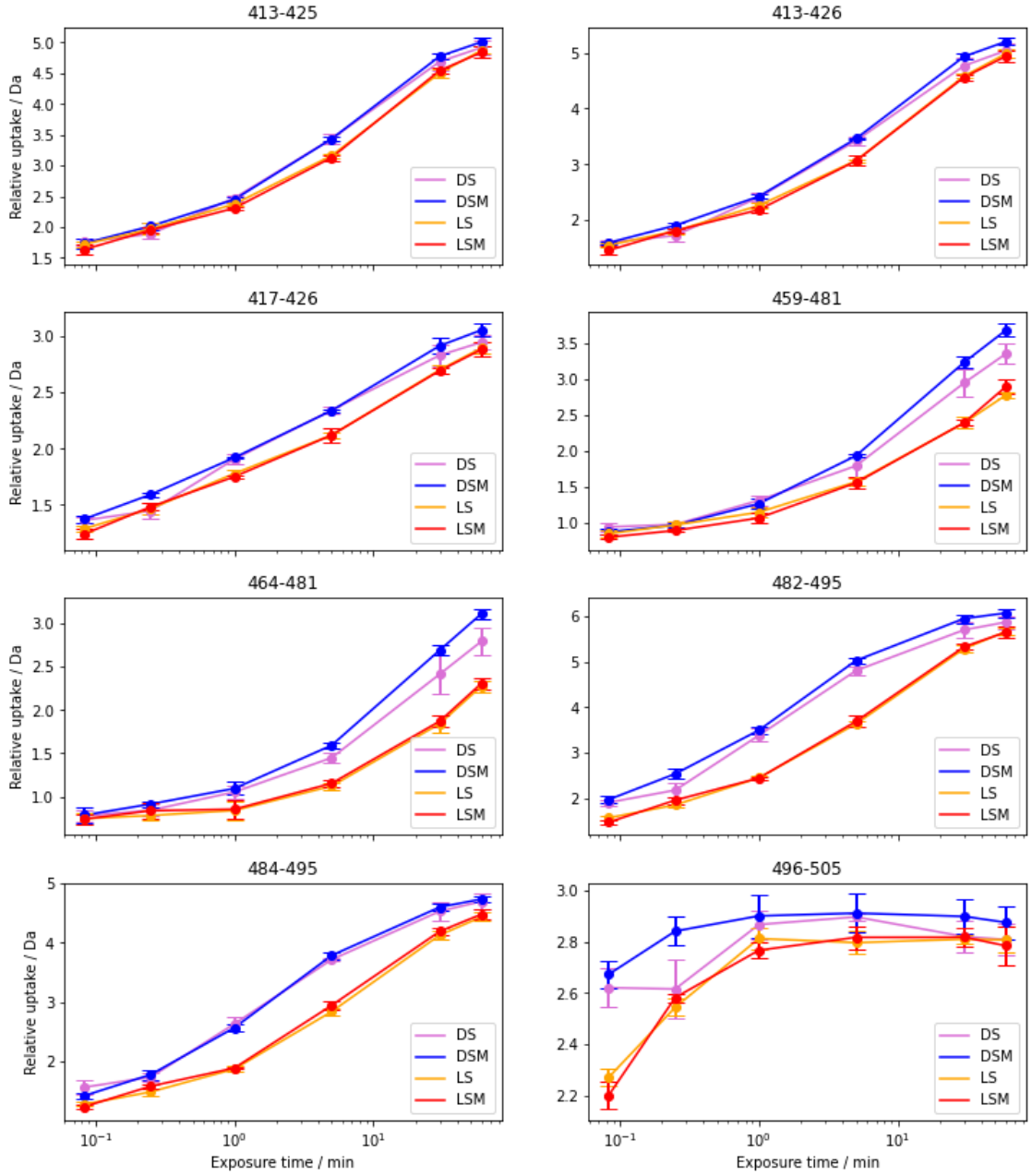

Figure A.9: Relative uptake of *Er*Cry4a W369F peptides of interest. Relative uptake of *Er*Cry4a W369F peptides that experience light-induced conformational changes over time displayed on a logarithmic time scale. For more information see chapter 3. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

A.2 Appendix for chapter 4

A.2.1 Relative fractional uptake

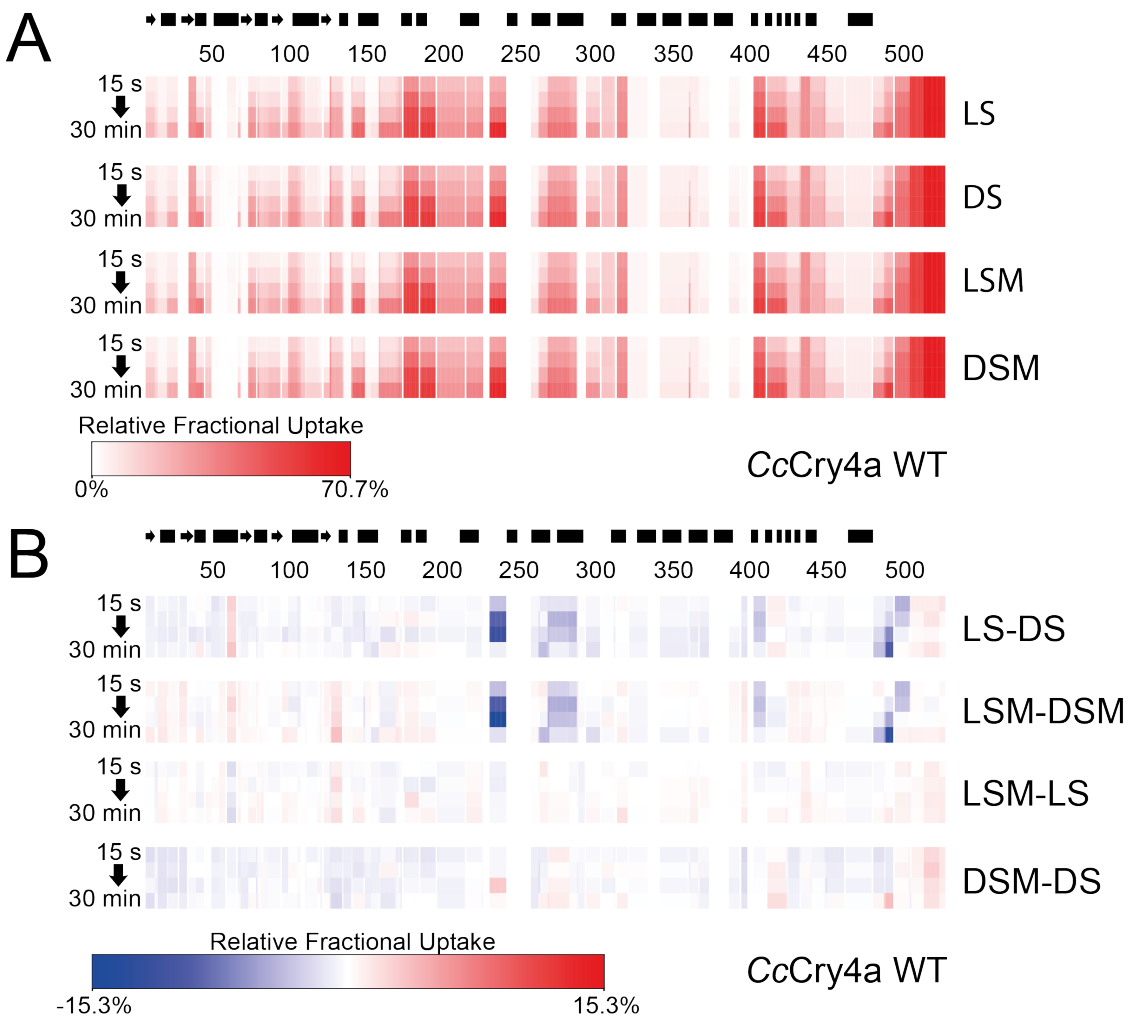

Figure A.10: Relative fractional uptake of *Cc*Cry4a WT. (**A**) Relative fractional uptake across the entire sequence measured at the peptide level. (**B**) Relative fractional uptake differences between states as indicated. As per usual convention, the N-terminus is situated on the left side and the C-terminus on the right. The time points displayed for both (A) and (B) are 15 s, 1 min, 5 min, and 30 min. For more detailed information please see Figure 4.4 in the main text. Beta sheets (arrows) and alpha-helices (rectangles) are indicated above the sequences [1]. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

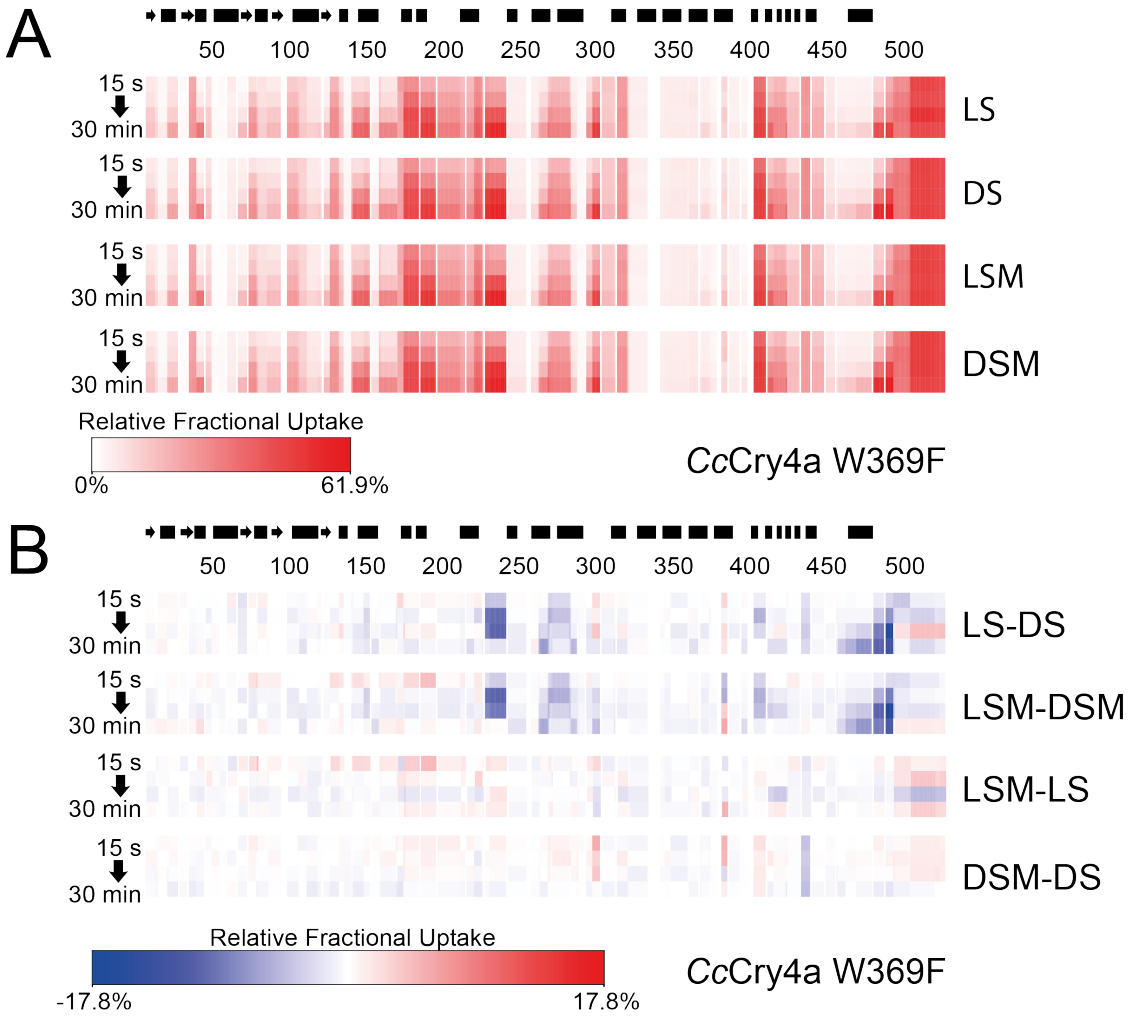

Figure A.11: Relative fractional uptake of *Cc*Cry4a W369F. (**A**) Relative fractional uptake across the entire sequence measured at the peptide level. (**B**) Relative fractional uptake differences between states as indicated. As per usual convention, the N-terminus is situated on the left side and the C-terminus on the right. The time points displayed for both (A) and (B) are 15 s, 1 min, 5 min, and 30 min. For more detailed information please see Figure 4.4 in the main text. Beta sheets (arrows) and alpha-helices (rectangles) are indicated above the sequences [1]. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

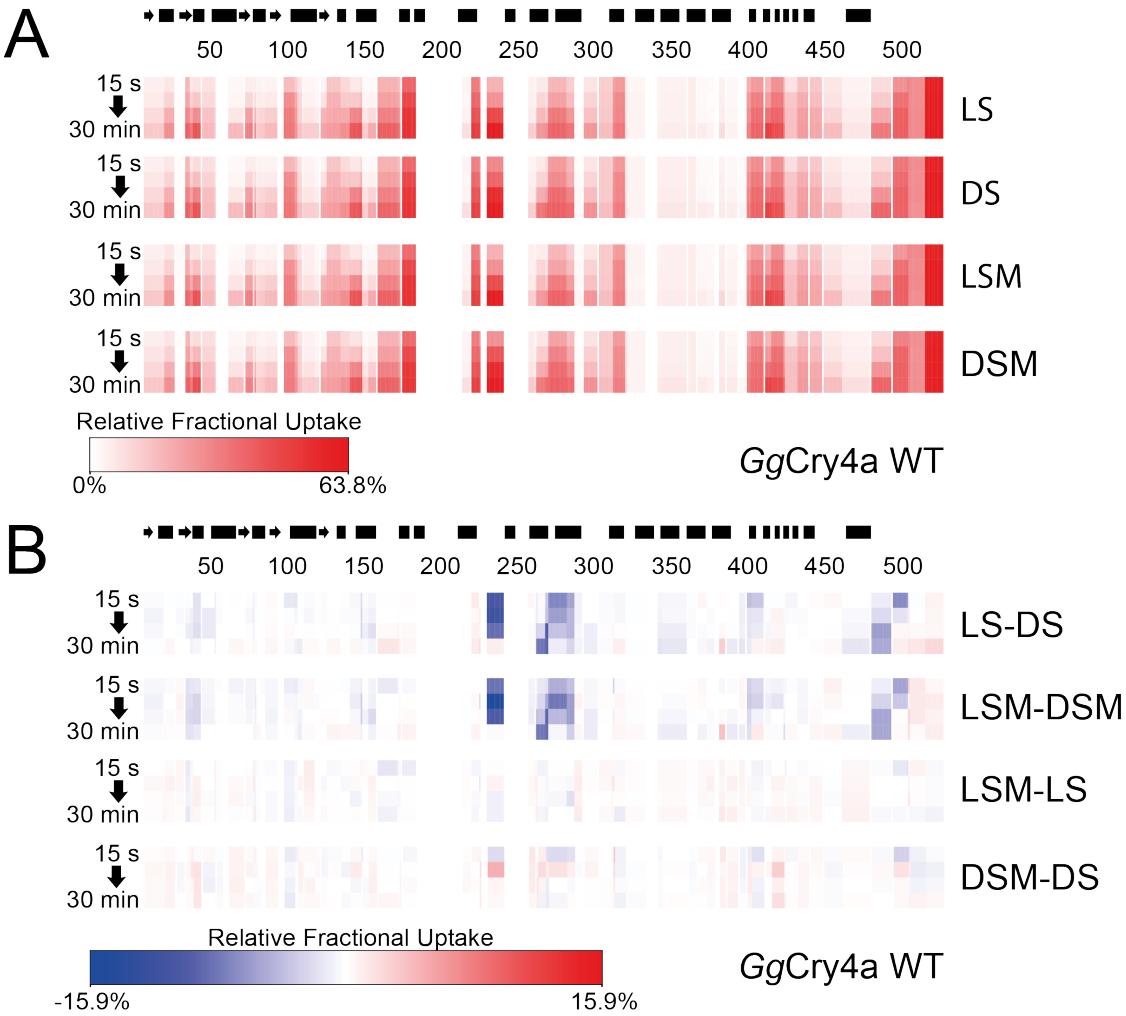

Figure A.12: Relative fractional uptake of *Gg*Cry4a WT. (**A**) Relative fractional uptake across the entire sequence measured at the peptide level. (**B**) Relative fractional uptake differences between states as indicated. As per usual convention, the N-terminus is situated on the left side and the C-terminus on the right. The time points displayed for both (A) and (B) are 15 s, 1 min, 5 min, and 30 min. For more detailed information please see Figure 4.4 in the main text. Beta sheets (arrows) and alpha-helices (rectangles) are indicated above the sequences [1]. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

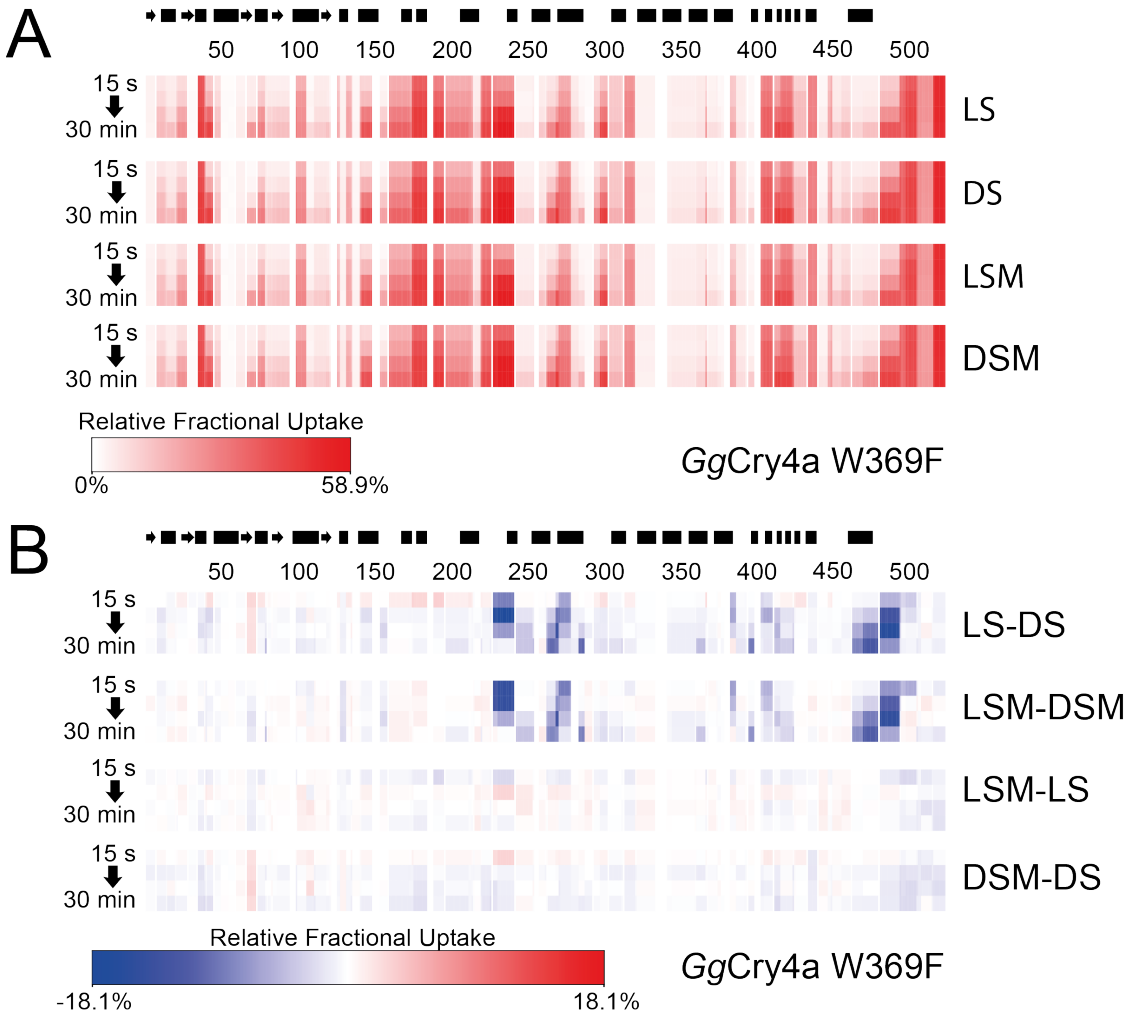

Figure A.13: Relative fractional uptake of *Gg*Cry4a W369F. (**A**) Relative fractional uptake across the entire sequence measured at the peptide level. (**B**) Relative fractional uptake differences between states as indicated. As per usual convention, the N-terminus is situated on the left side and the C-terminus on the right. The time points displayed for both (A) and (B) are 15 s, 1 min, 5 min, and 30 min. For more detailed information please see Figure 4.4 in the main text. Beta sheets (arrows) and alpha-helices (rectangles) are indicated above the sequences [1]. DS - dark state, DSM - dark state + magnet, LS - light state, LSM - light state + magnet.

A.2.2 Chiclet plots created in HD-eXplosion

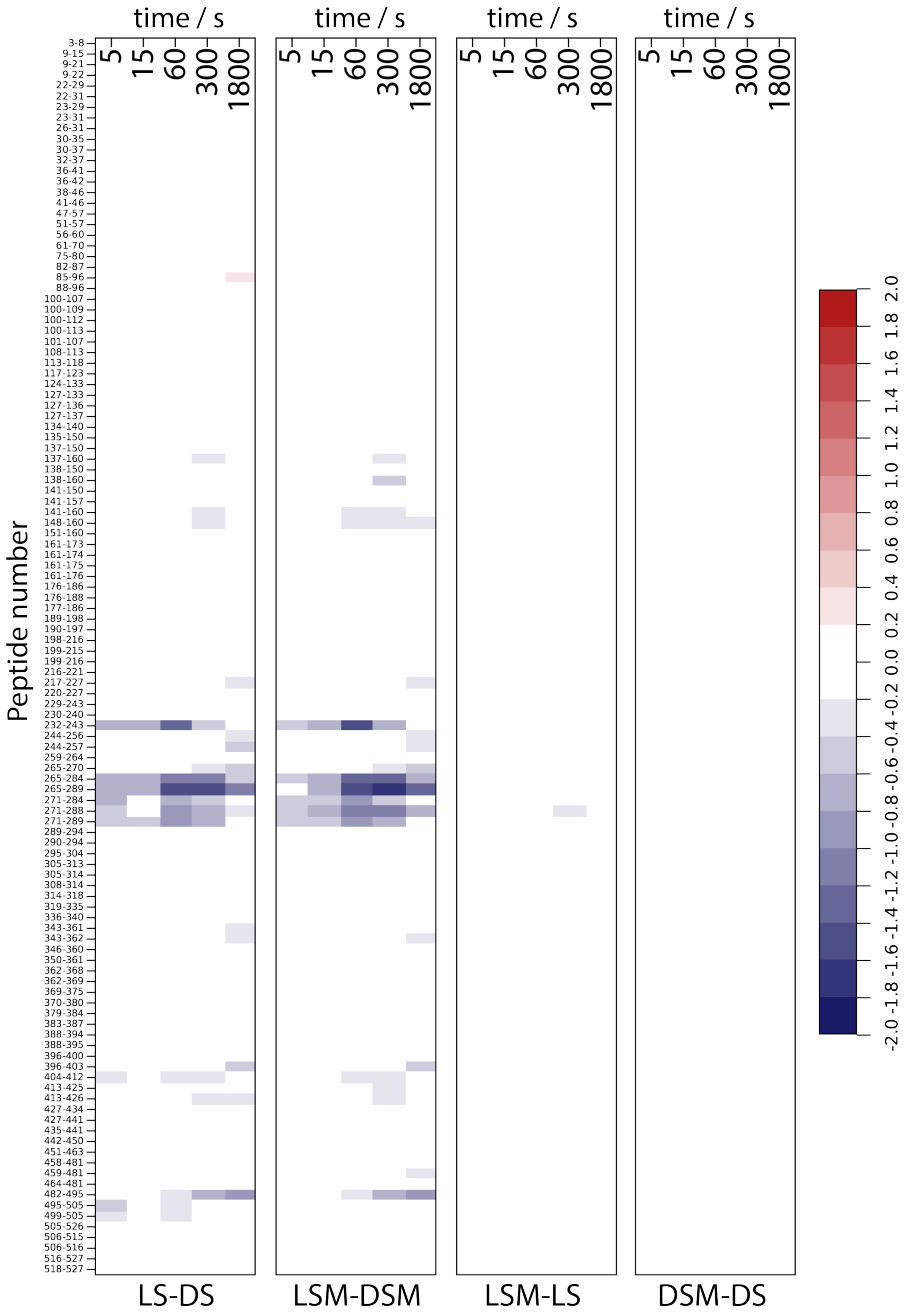

Figure A.14: HD-eXplosion chiclet plots of *Er*Cry4a WT significant peptides. For details please see section 4.1 in chapter 4.

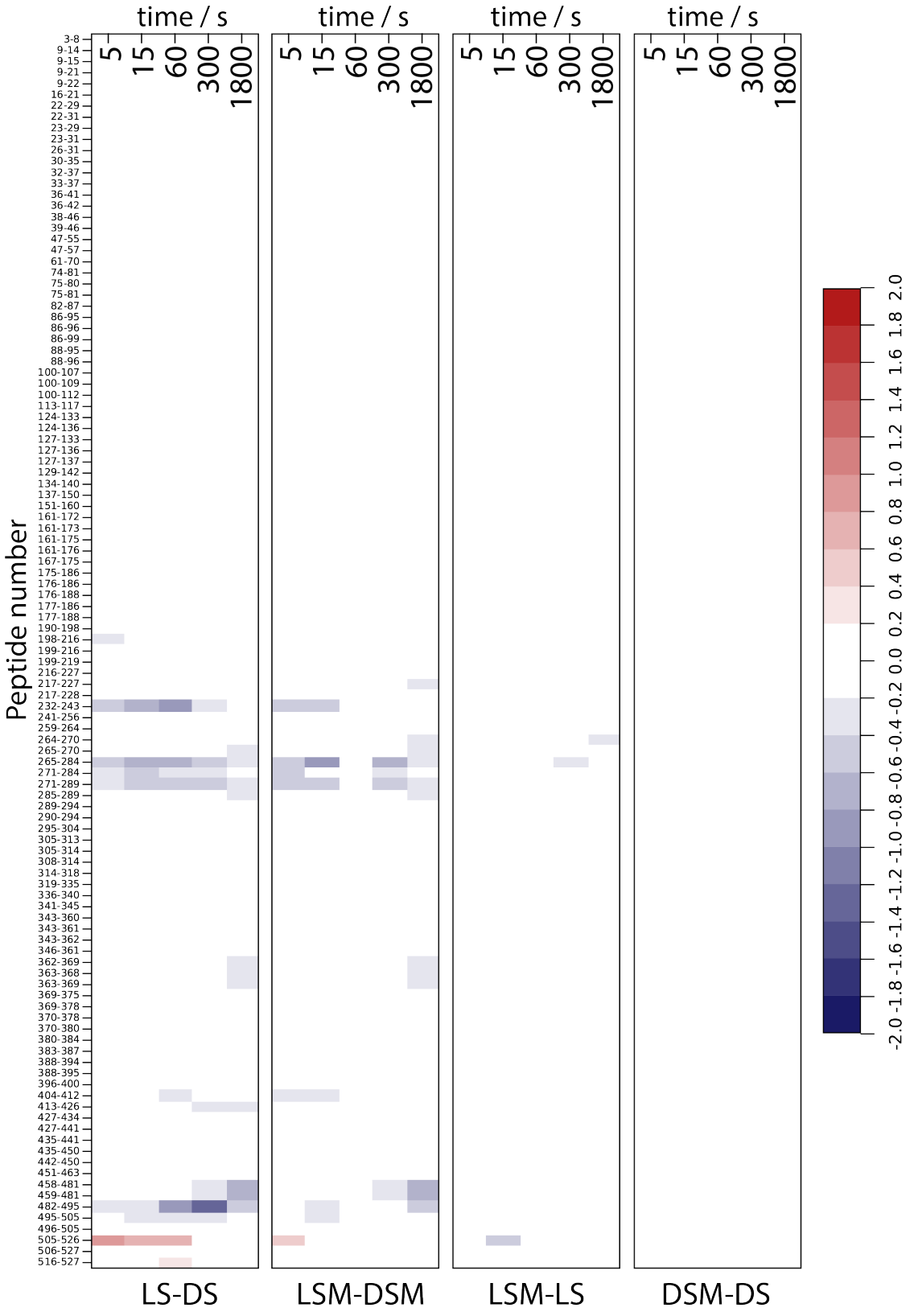

Figure A.15: HD-eXplosion chiclet plots of *Er*Cry4a W369F significant peptides. For details please see section 4.1 in chapter 4.

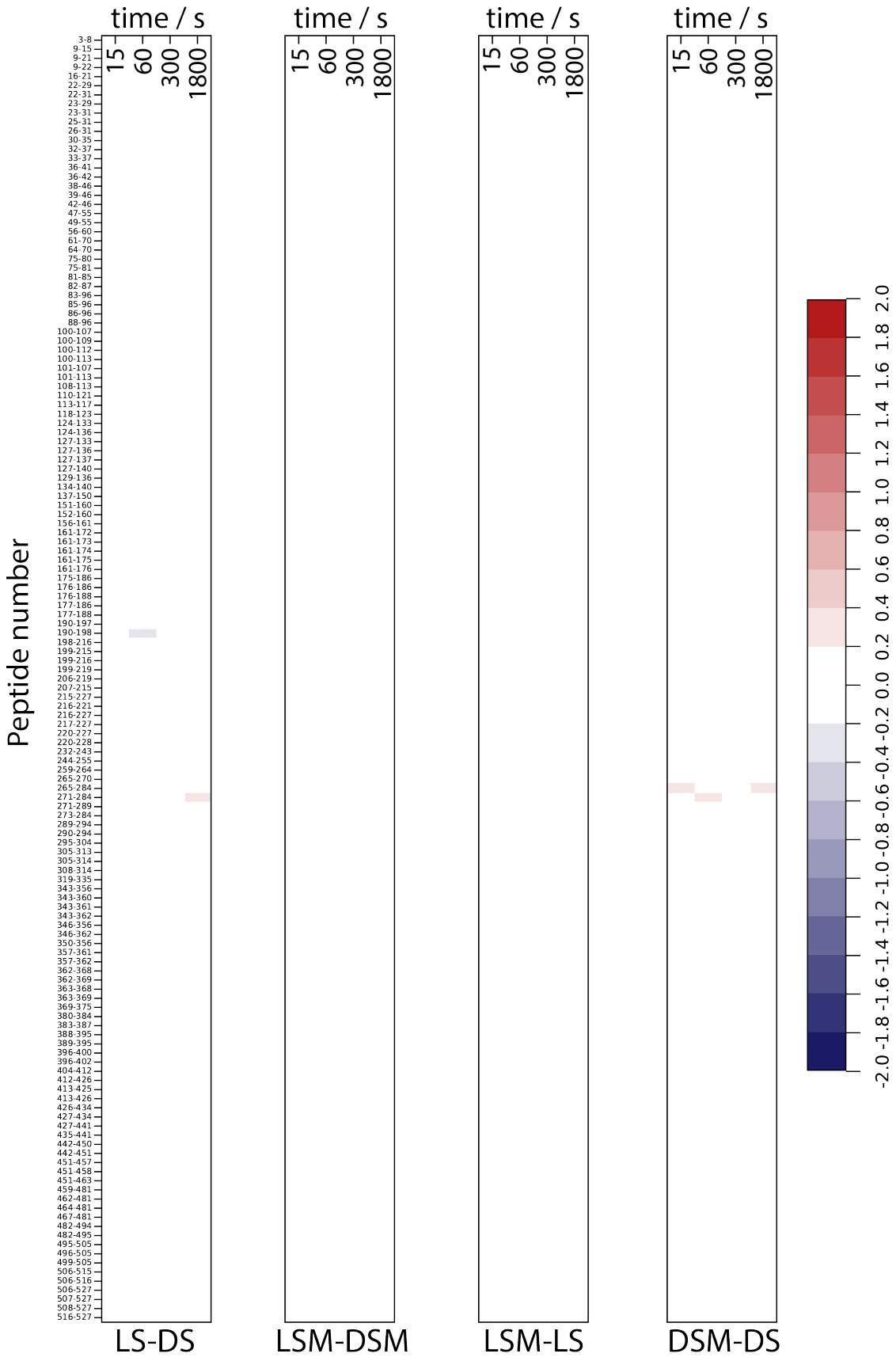

Figure A.16: HD-eXplosion chiclet plots of *Er*Cry4a R356E significant peptides. For details please see section 4.1 in chapter 4.

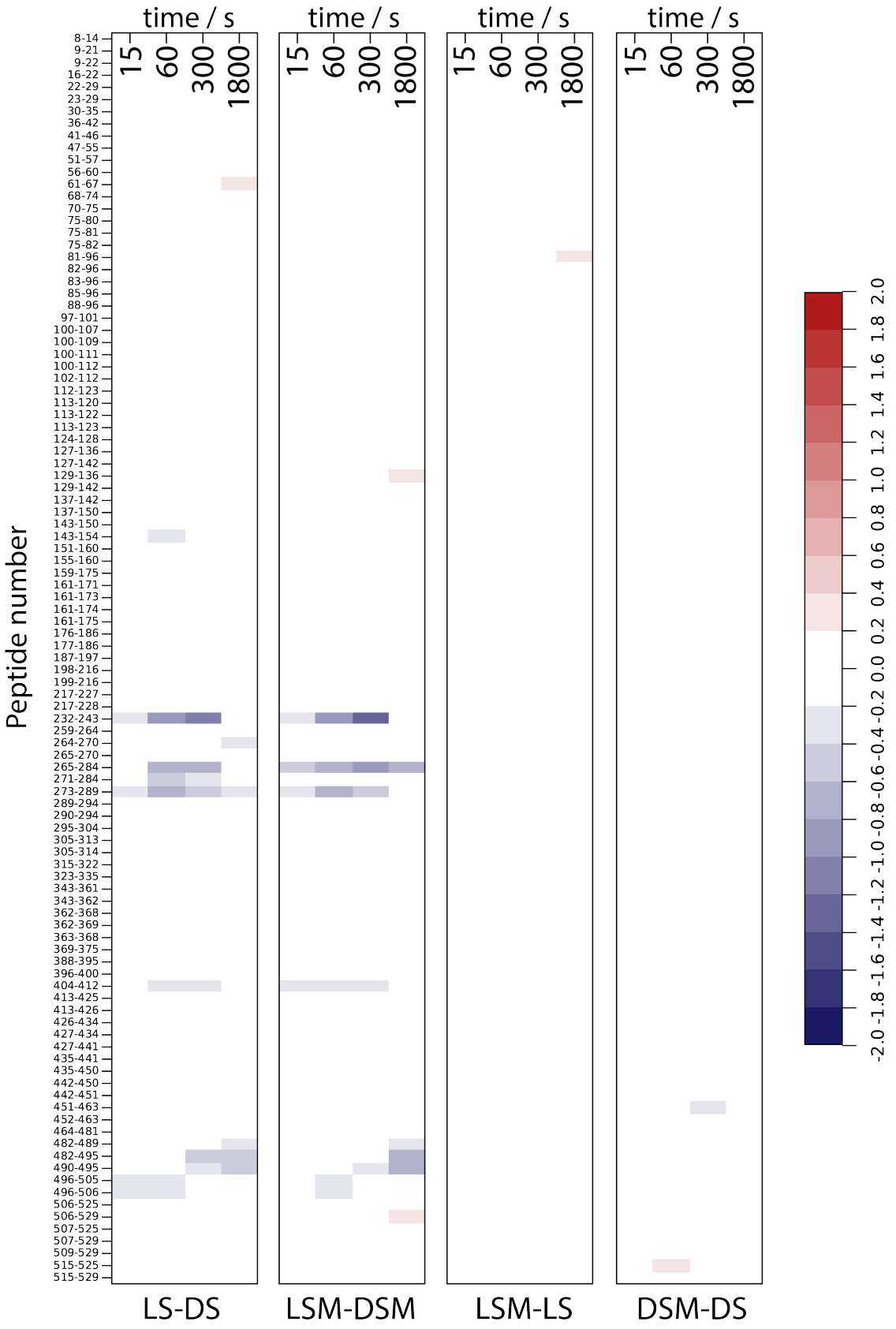

Figure A.17: HD-eXplosion chiclet plots of *Cc*Cry4a WT significant peptides. For details please see section 4.1 in chapter 4.

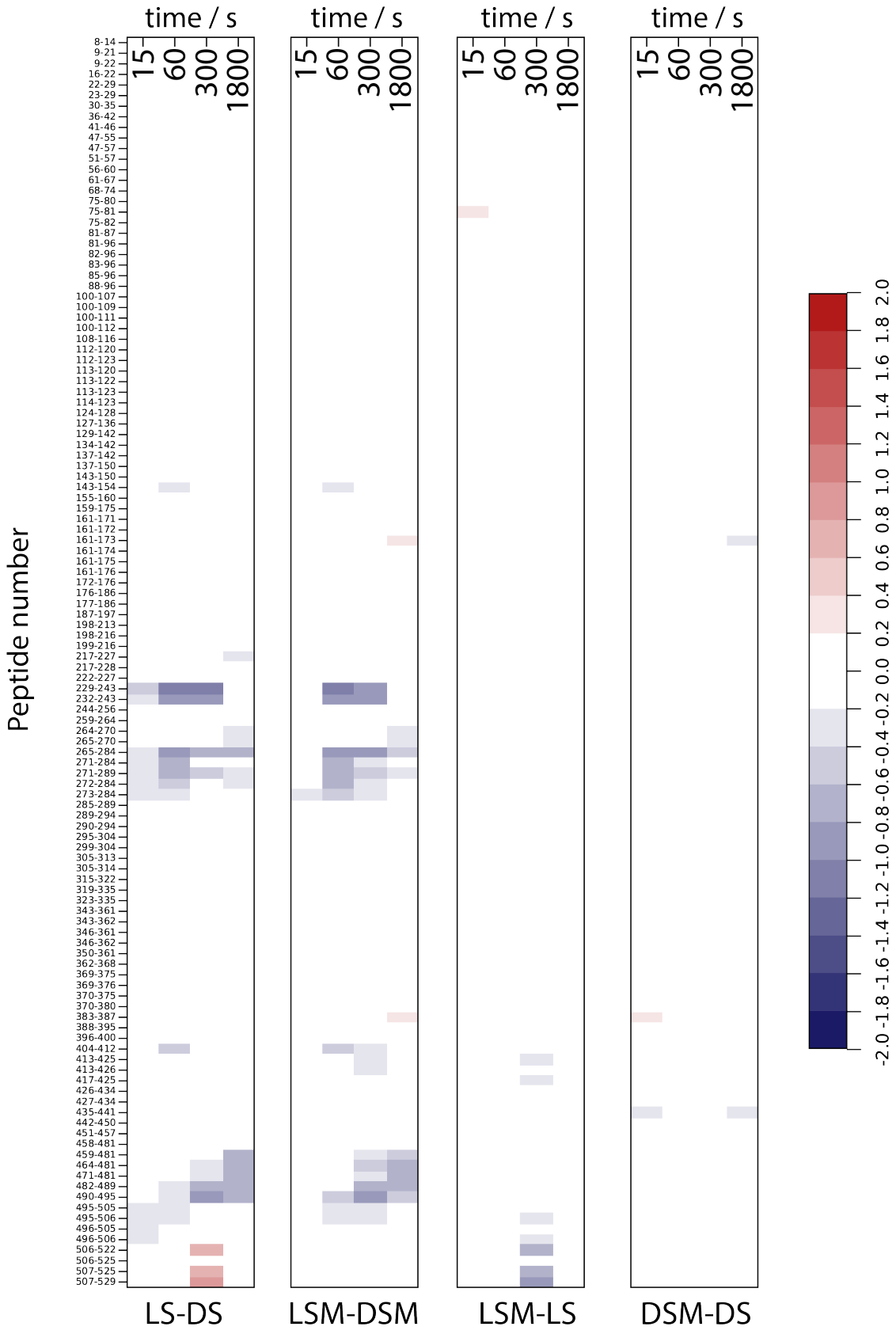

Figure A.18: HD-eXplosion chiclet plots of *Cc*Cry4a W369F significant peptides. For details please see section 4.1 in chapter 4.

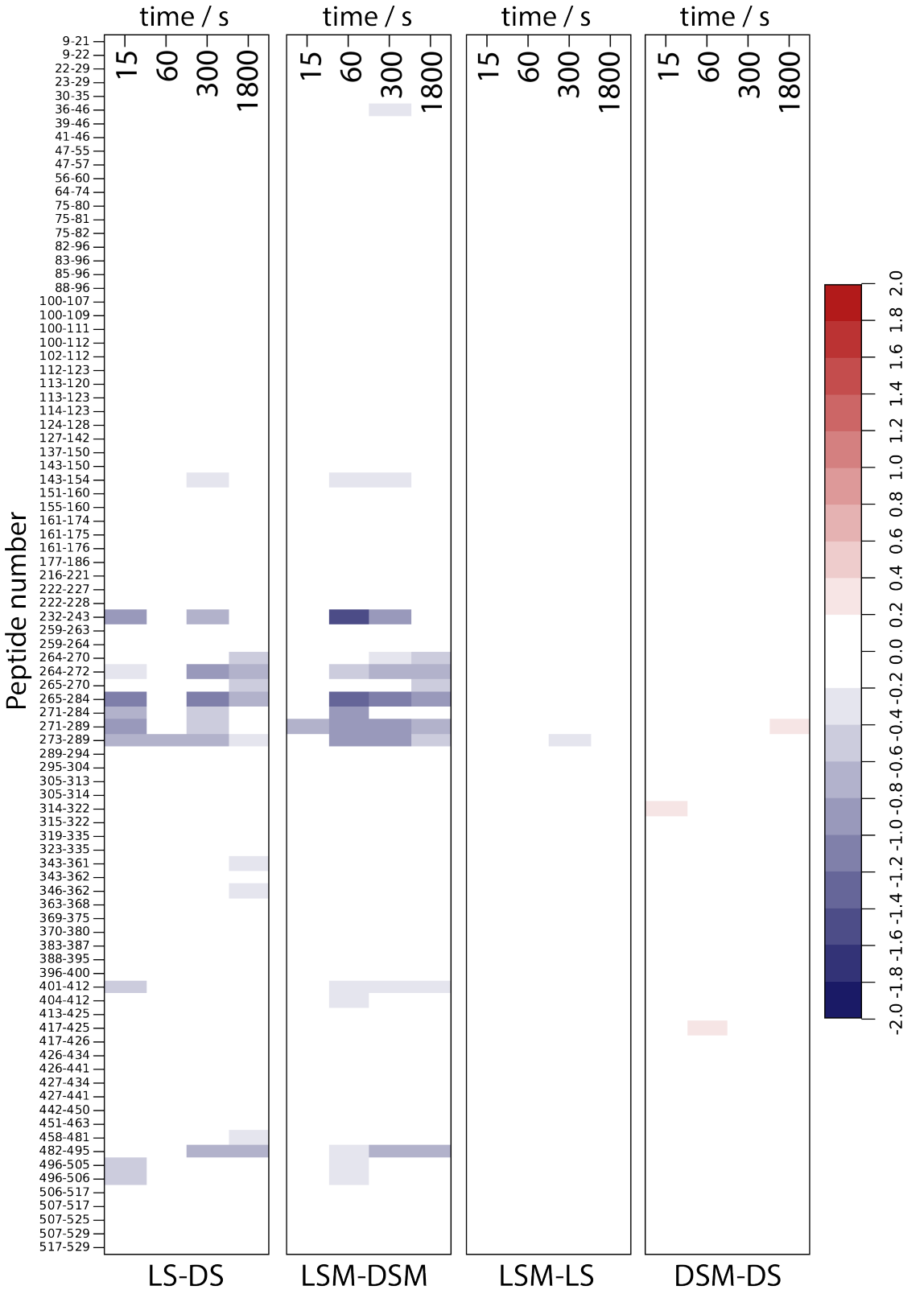

Figure A.19: HD-eXplosion chiclet plots of *Gg*Cry4a WT significant peptides. For details please see section 4.1 in chapter 4.

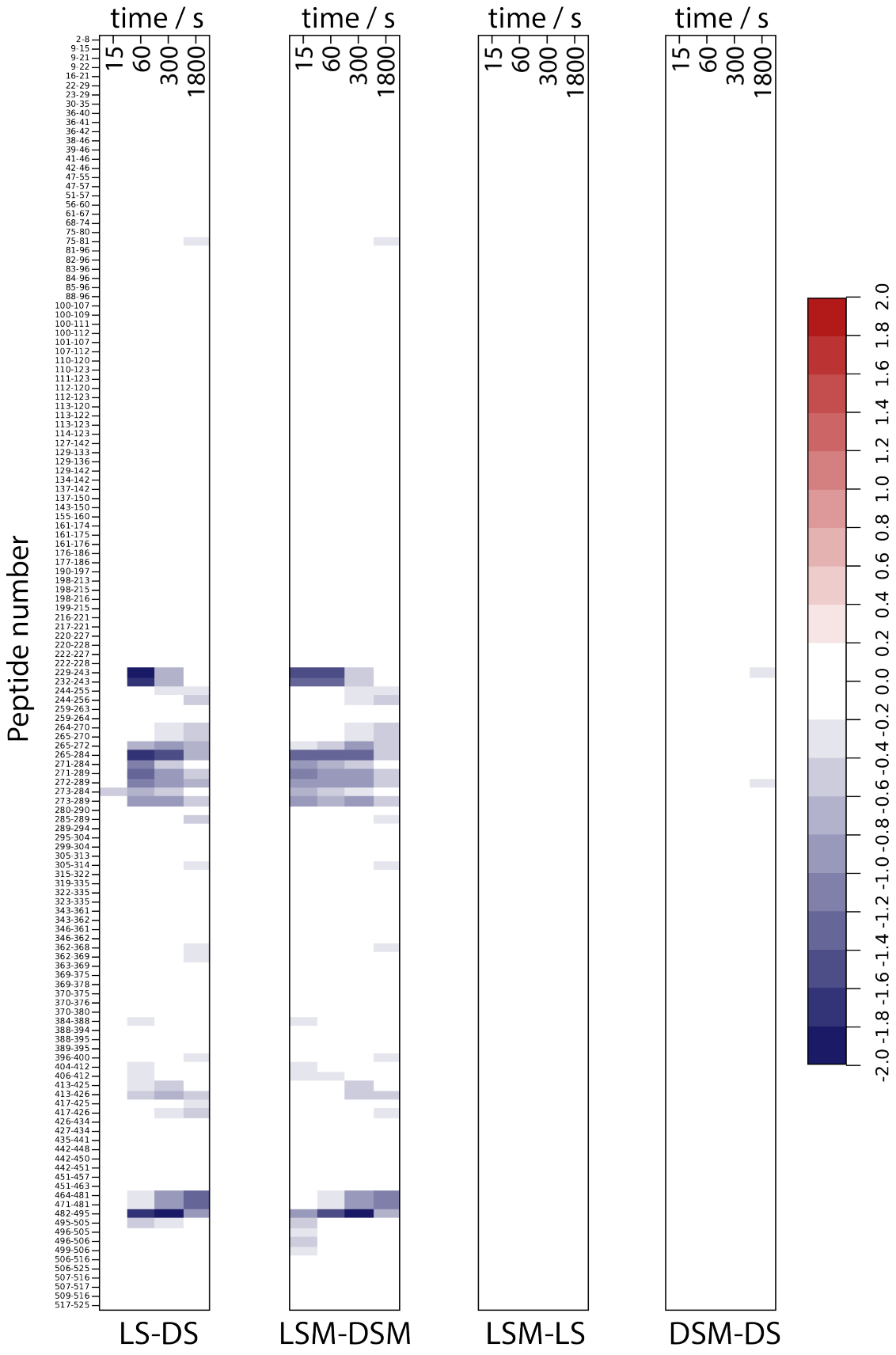

Figure A.20: HD-eXplosion chiclet plots of *Gg*Cry4a W369F significant peptides. For details please see section 4.1 in chapter 4.

A.2.3 ΔRFU plots created by Laila M. N. Shah

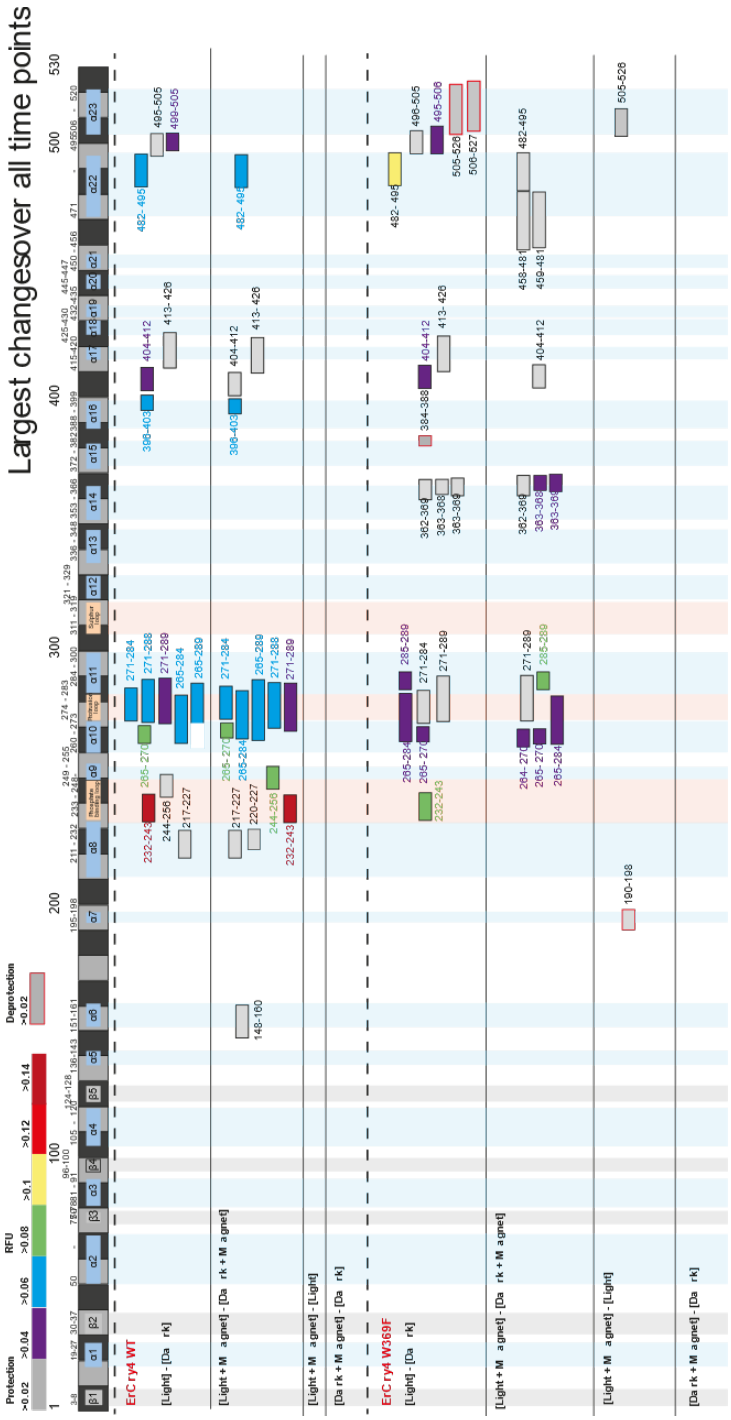

Figure A.21: Relative fractional uptake of *Er*Cry4a WT and W369F significant peptides. For details please see section 4.2.3 in chapter 4.

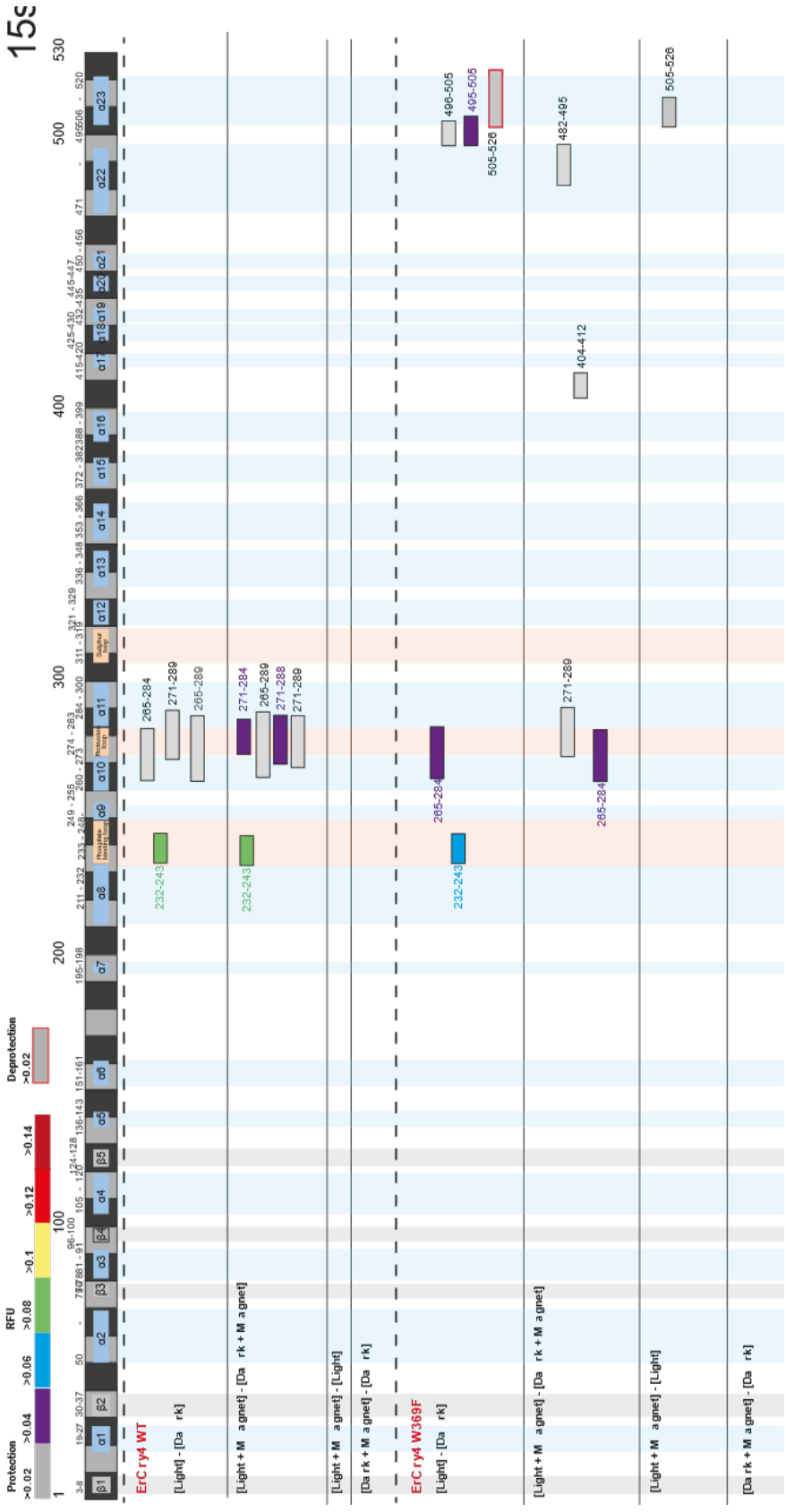

Figure A.22: Relative fractional uptake of *Er*Cry4a WT and W369F significant peptides after 15 s of labelling time. For details please see section 4.2.3 in chapter 4.

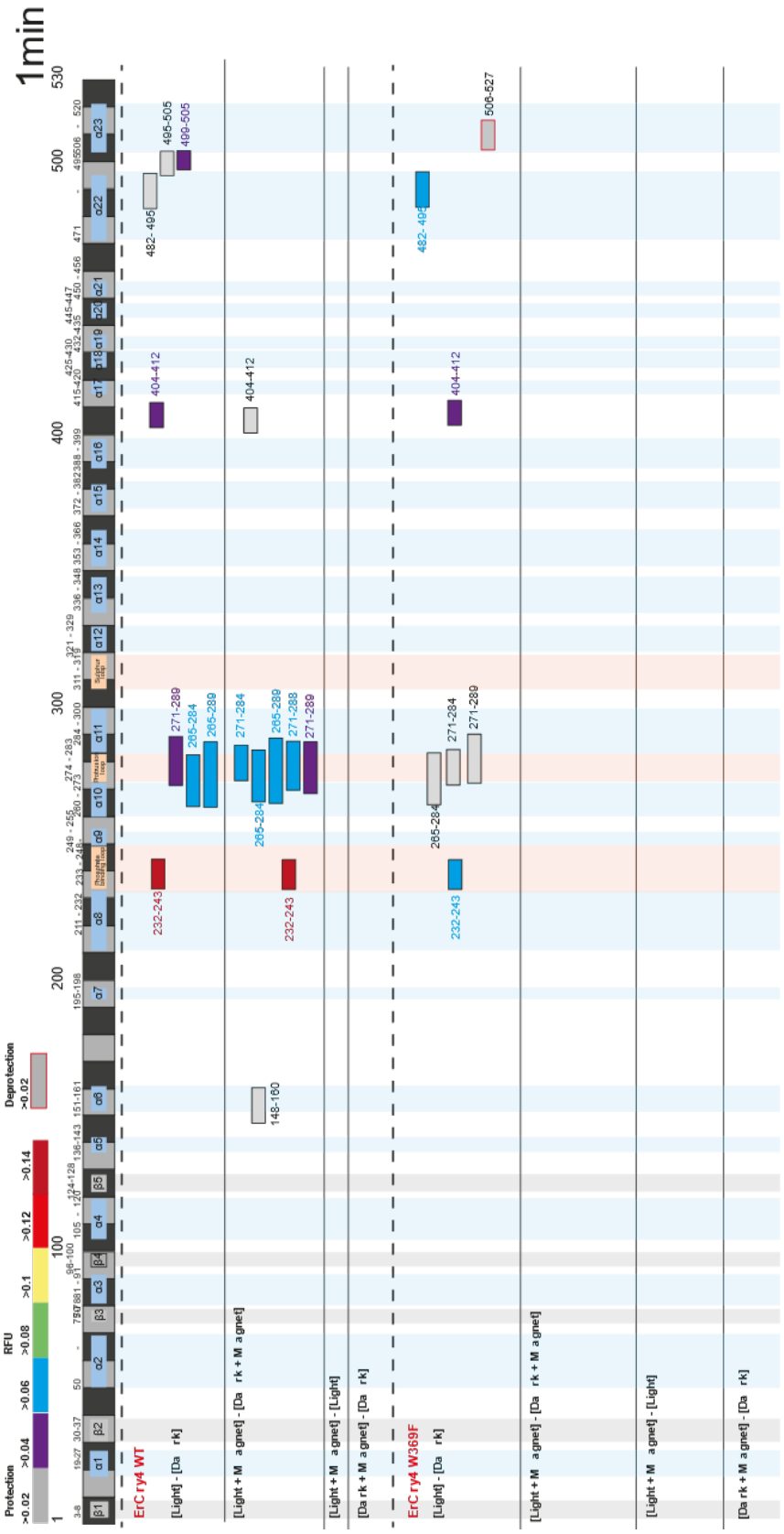

Figure A.23: Relative fractional uptake of *Er*Cry4a WT and W369F significant peptides after 1 min of labelling time. For details please see section 4.2.3 in chapter 4.

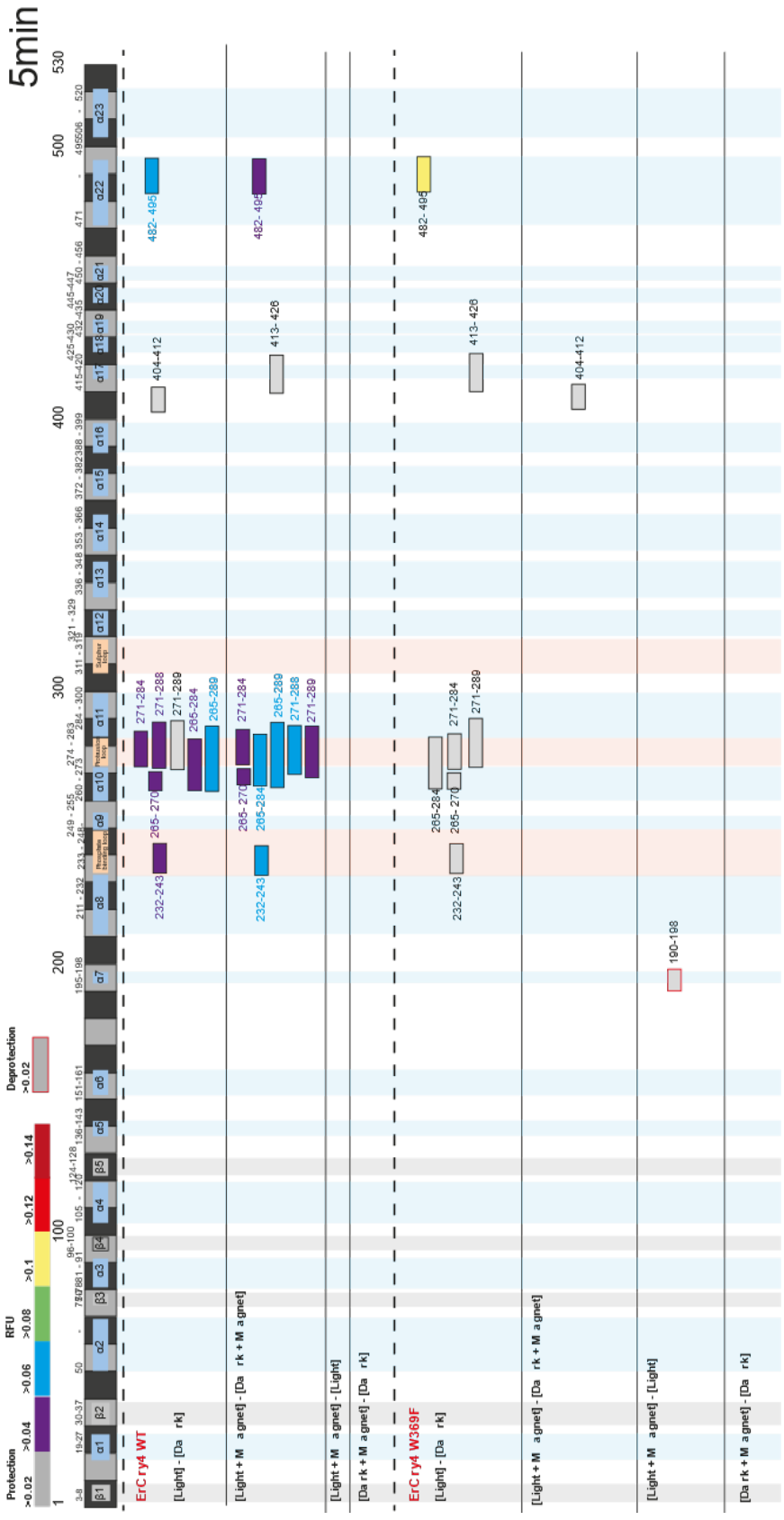

Figure A.24: Relative fractional uptake of *Er*Cry4a WT and W369F significant peptides after 5 min of labelling time. For details please see section 4.2.3 in chapter 4.

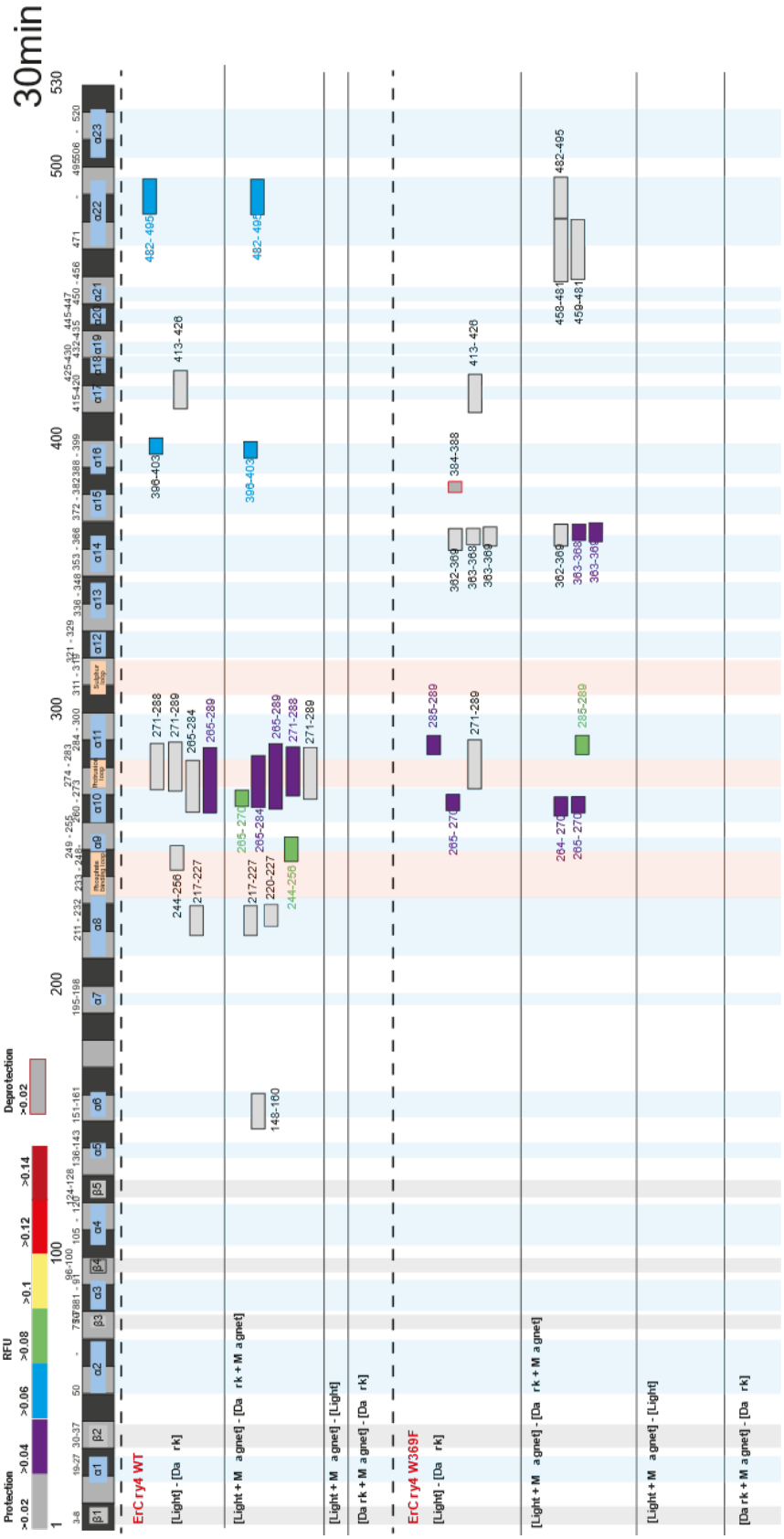

Figure A.25: Relative fractional uptake of *Er*Cry4a WT and W369F significant peptides after 30 min of labelling time. For details please see section 4.2.3 in chapter 4.

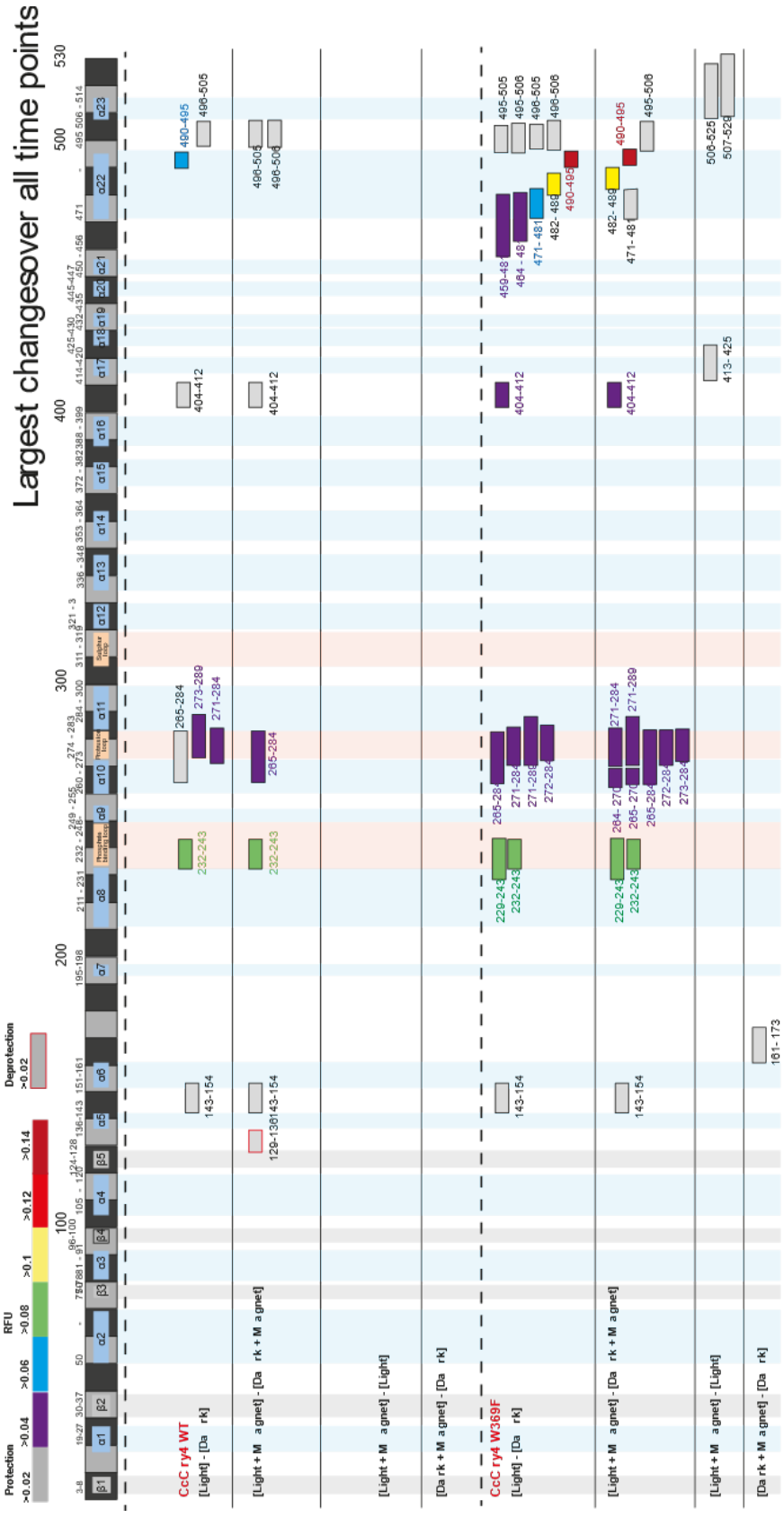

Figure A.26: Relative fractional uptake of *Cc*Cry4a WT and W369F significant peptides. For details please see section 4.2.3 in chapter 4.

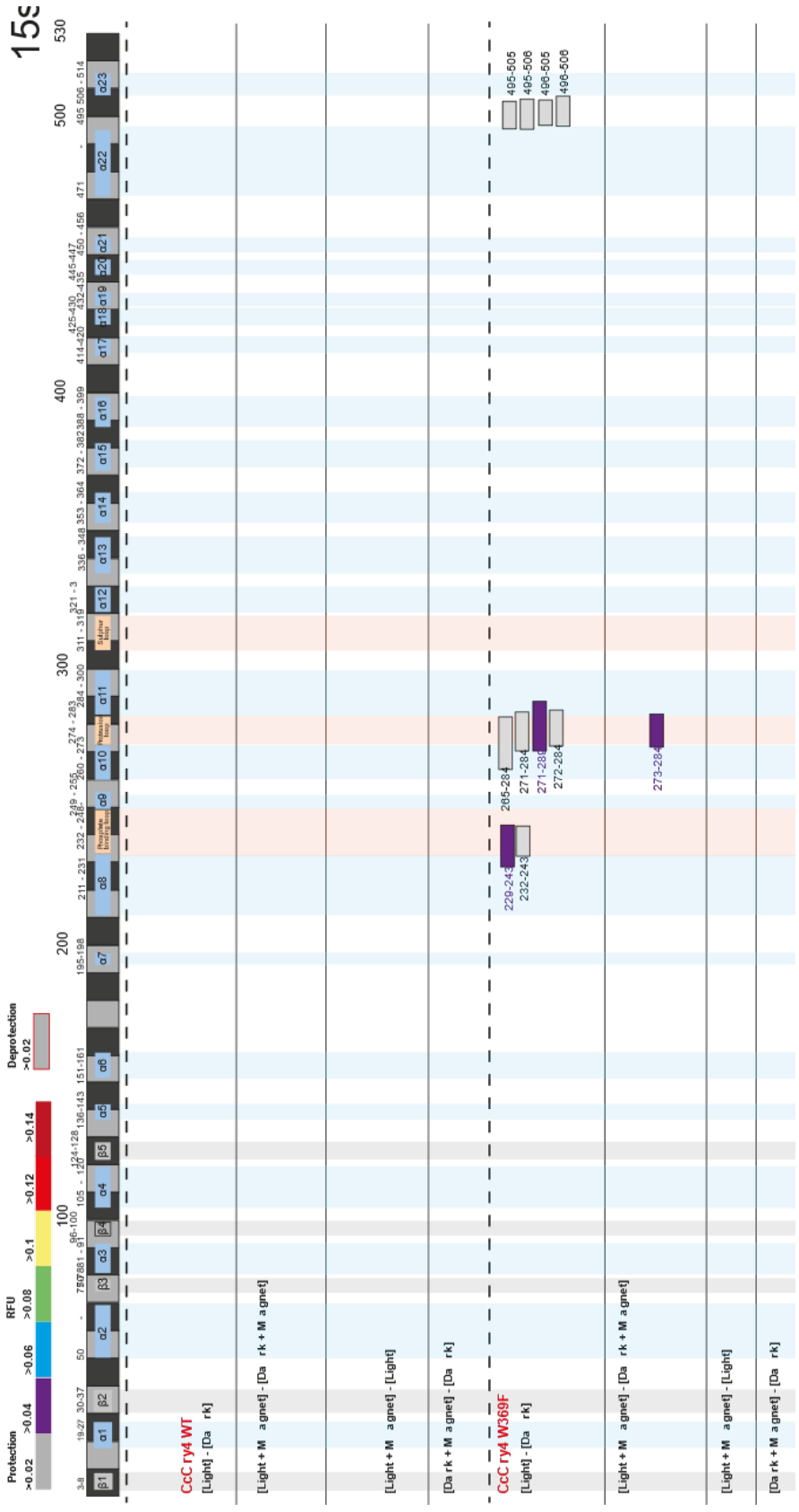

Figure A.27: Relative fractional uptake of *Cc*Cry4a WT and W369F significant peptides after 15 s of labelling time. For details please see section 4.2.3 in chapter 4.

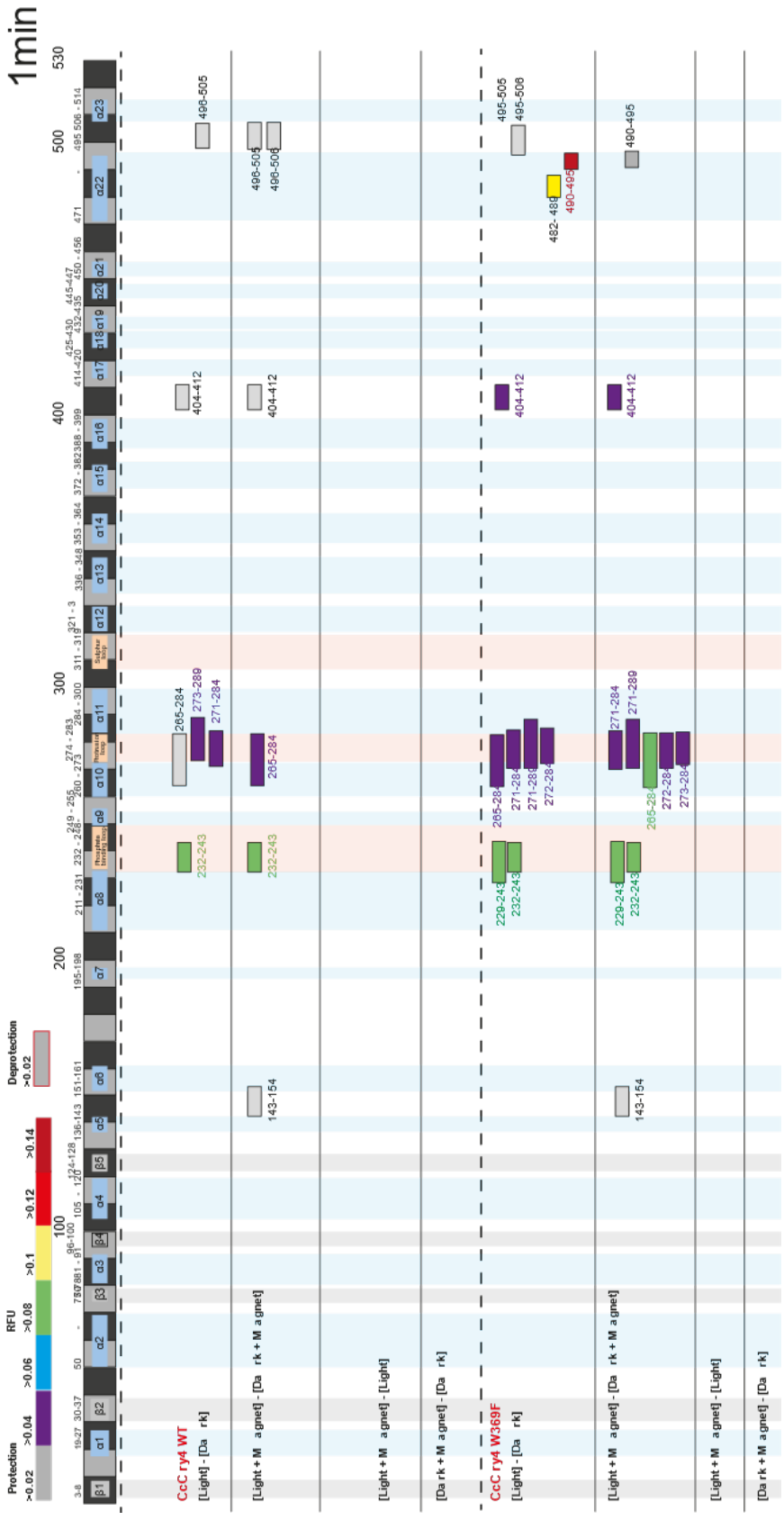

Figure A.28: Relative fractional uptake of *Cc*Cry4a WT and W369F significant peptides after 1 min of labelling time. For details please see section 4.2.3 in chapter 4.

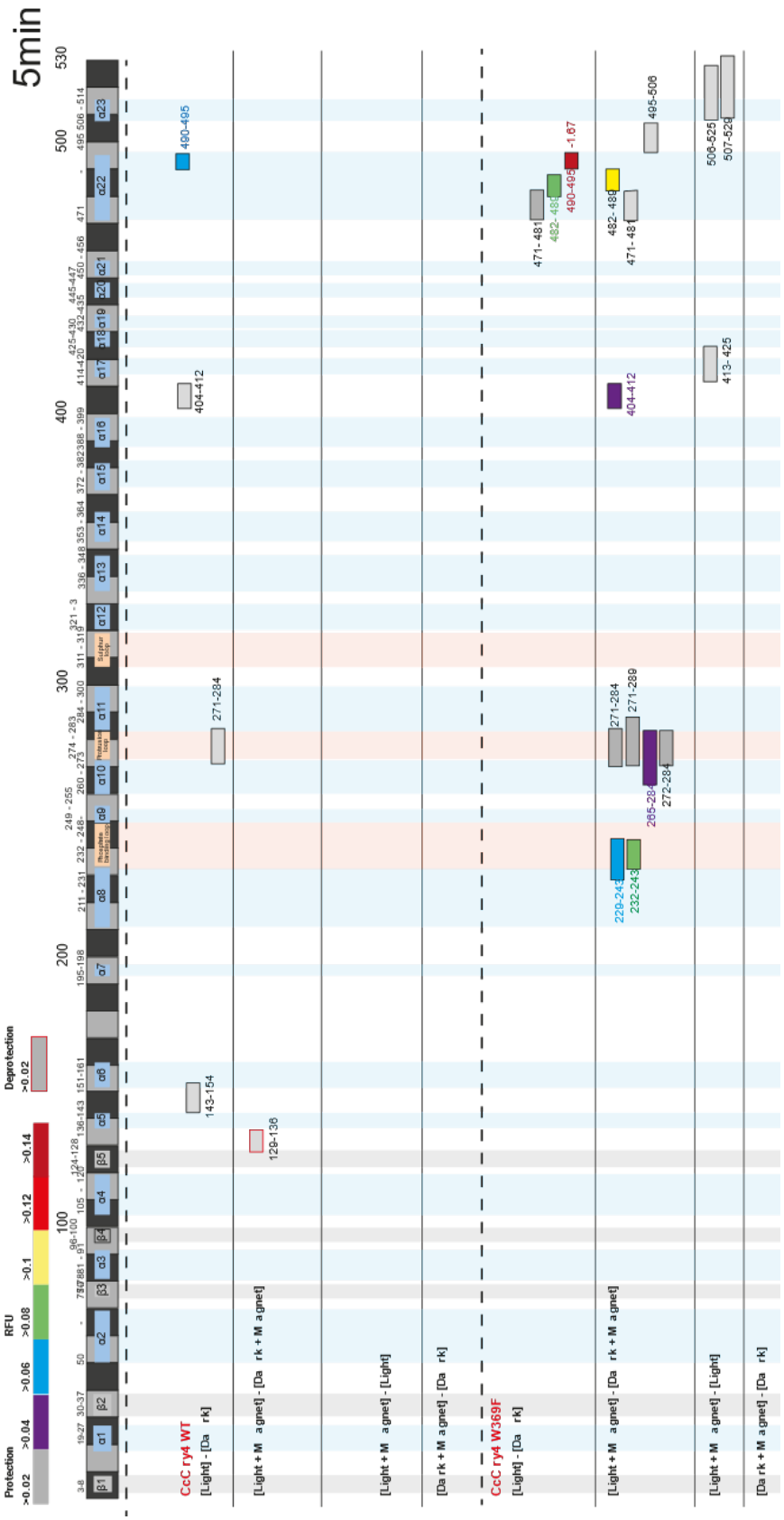

Figure A.29: Relative fractional uptake of *Cc*Cry4a WT and W369F significant peptides after 5 min of labelling time. For details please see section 4.2.3 in chapter 4.

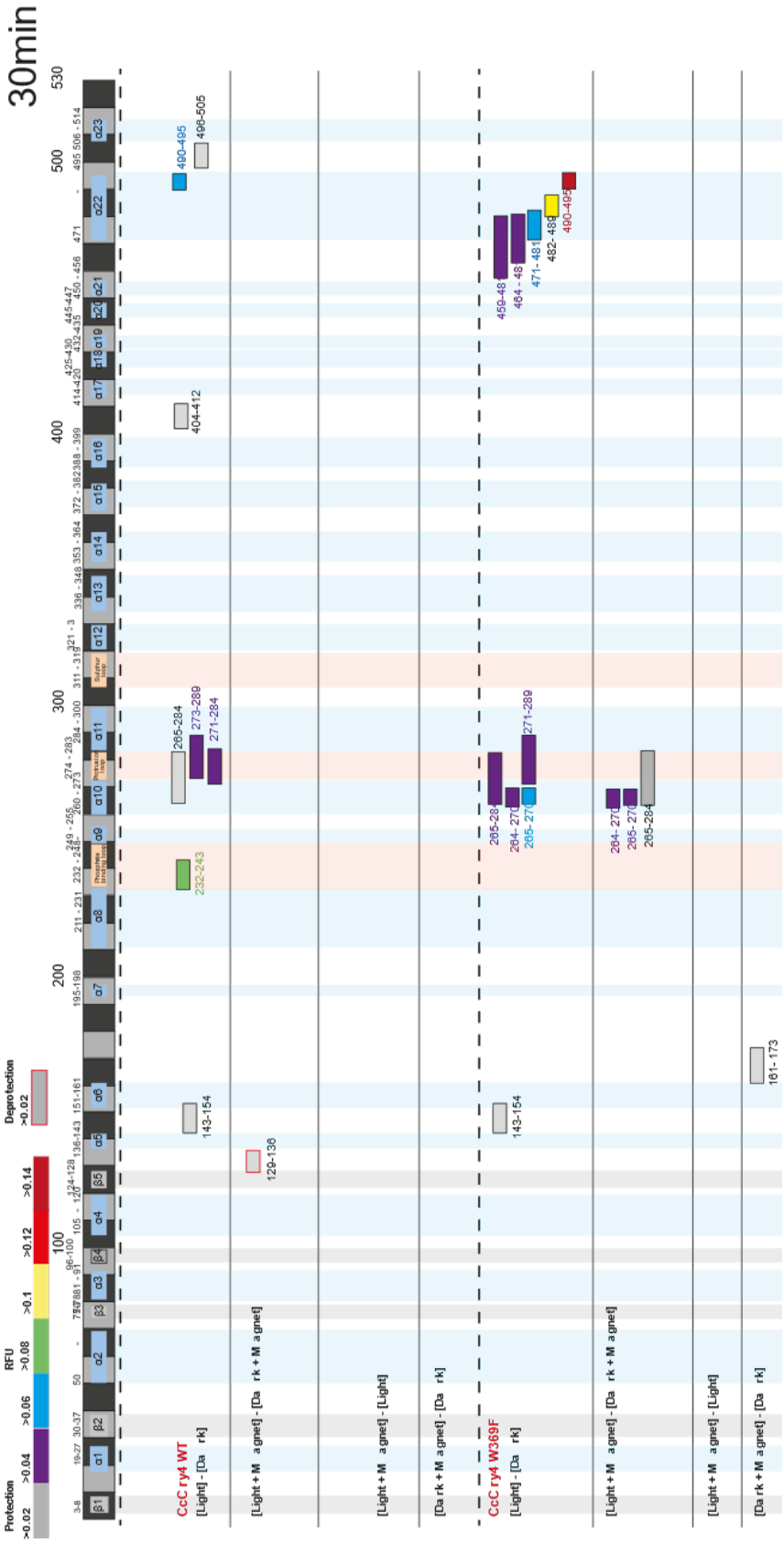

Figure A.30: Relative fractional uptake of *Cc*Cry4a WT and W369F significant peptides after 30 min of labelling time. For details please see section 4.2.3 in chapter 4.

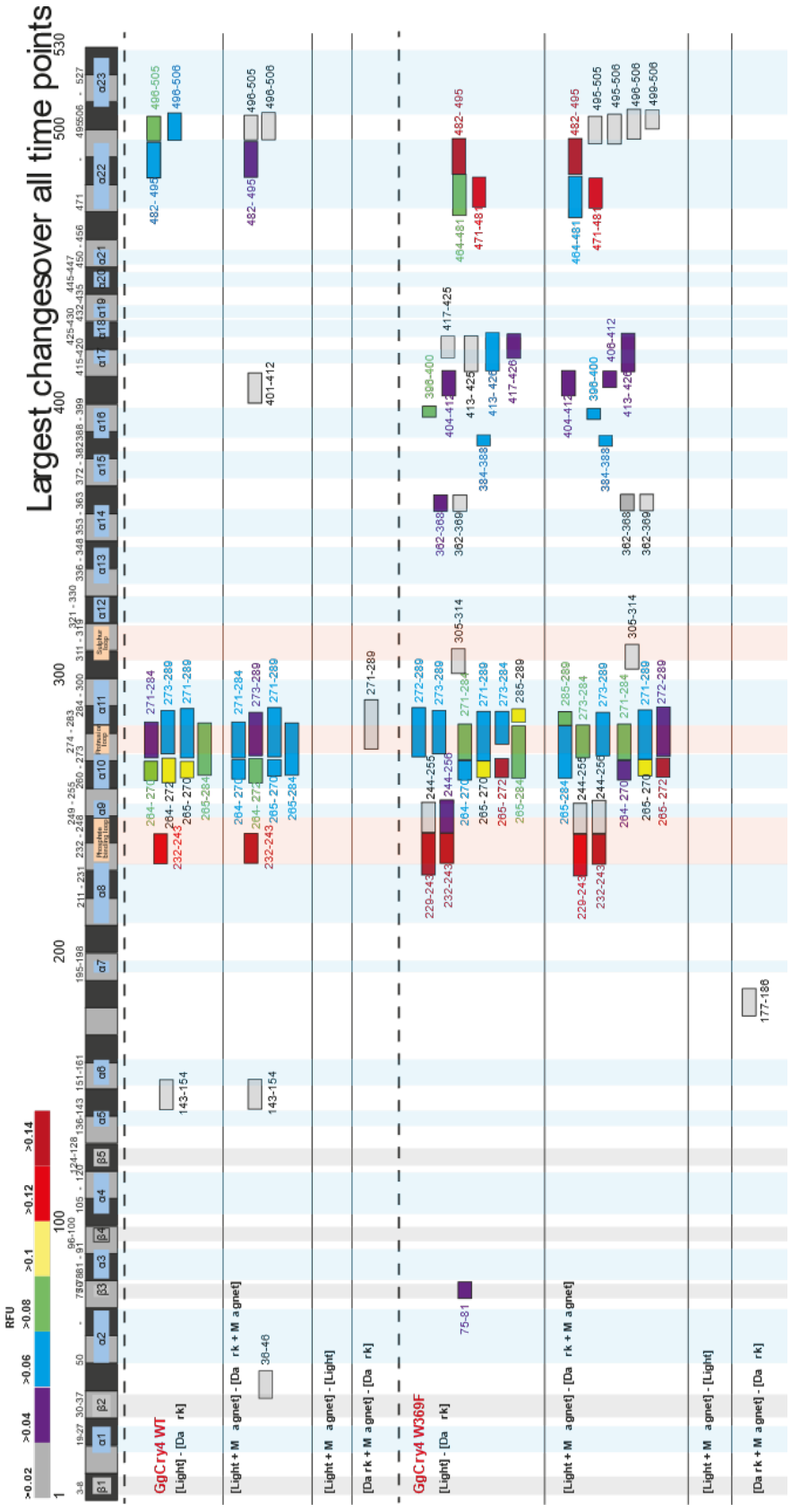

Figure A.31: Relative fractional uptake of *Gg*Cry4a WT and W369F significant peptides. For details please see section 4.2.3 in chapter 4.

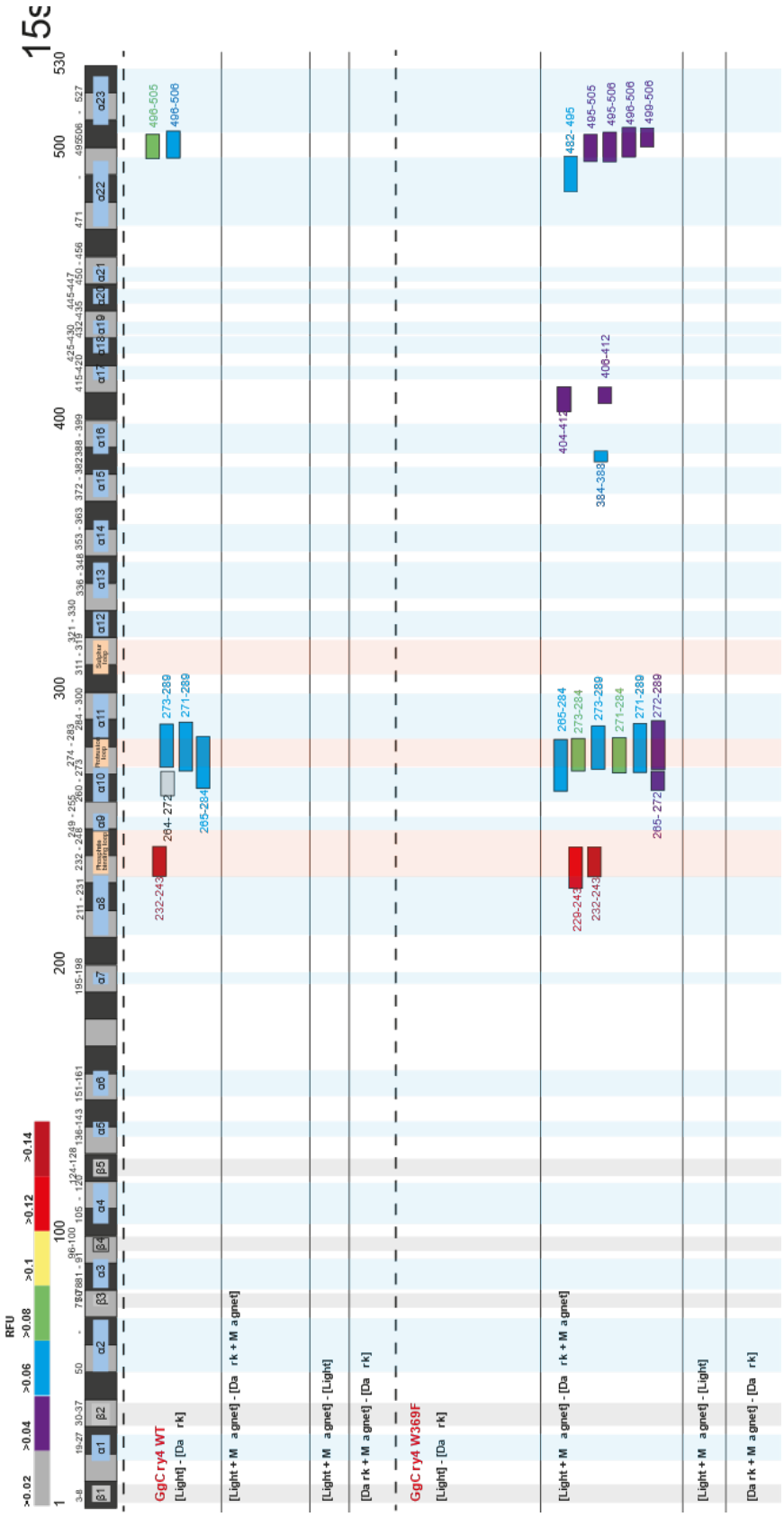

Figure A.32: Relative fractional uptake of *Gg*Cry4a WT and W369F significant peptides after 15 s of labelling time. For details please see section 4.2.3 in chapter 4.

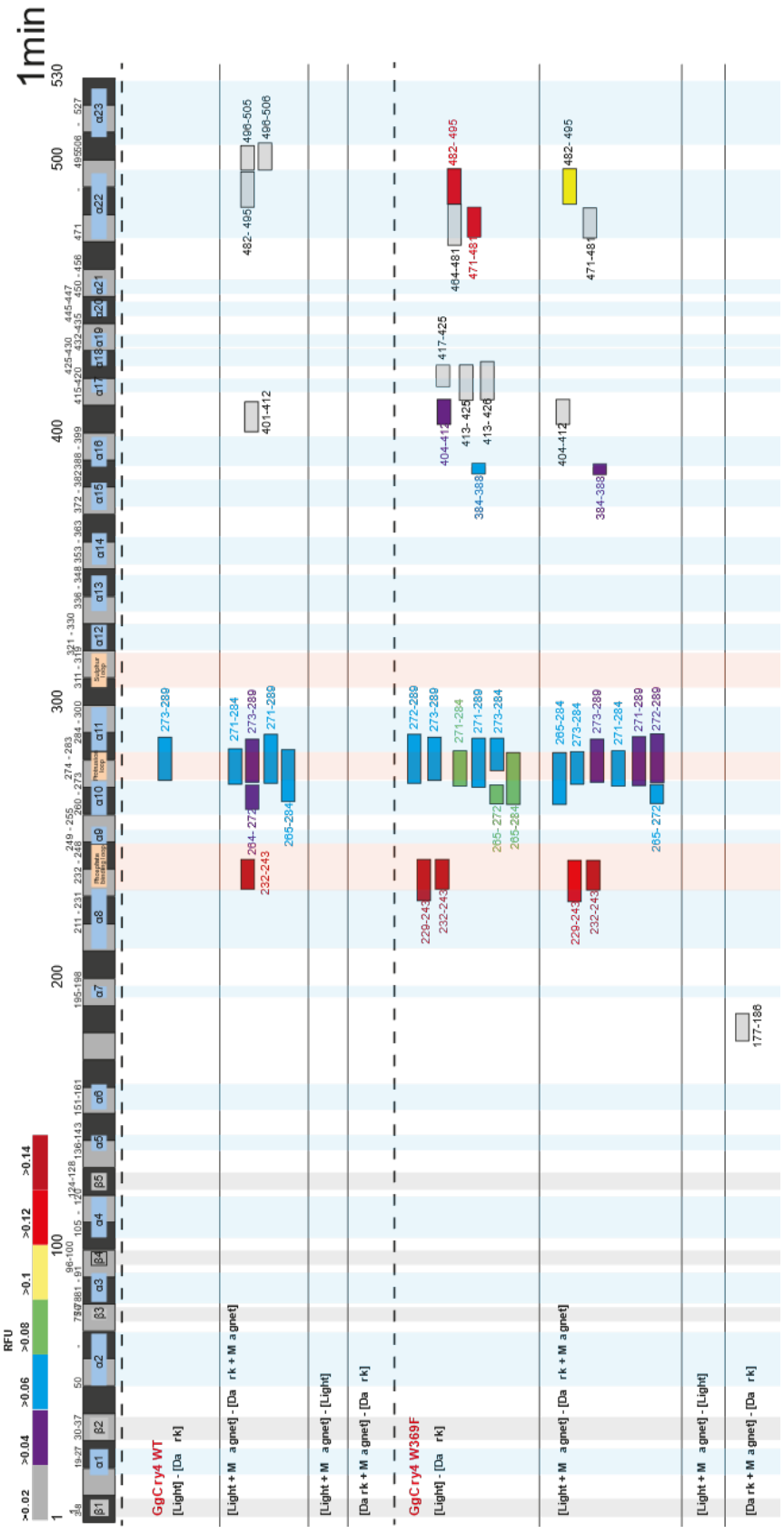

Figure A.33: Relative fractional uptake of *Gg*Cry4a WT and W369F significant peptides after 1 min of labelling time. For details please see section 4.2.3 in chapter 4.

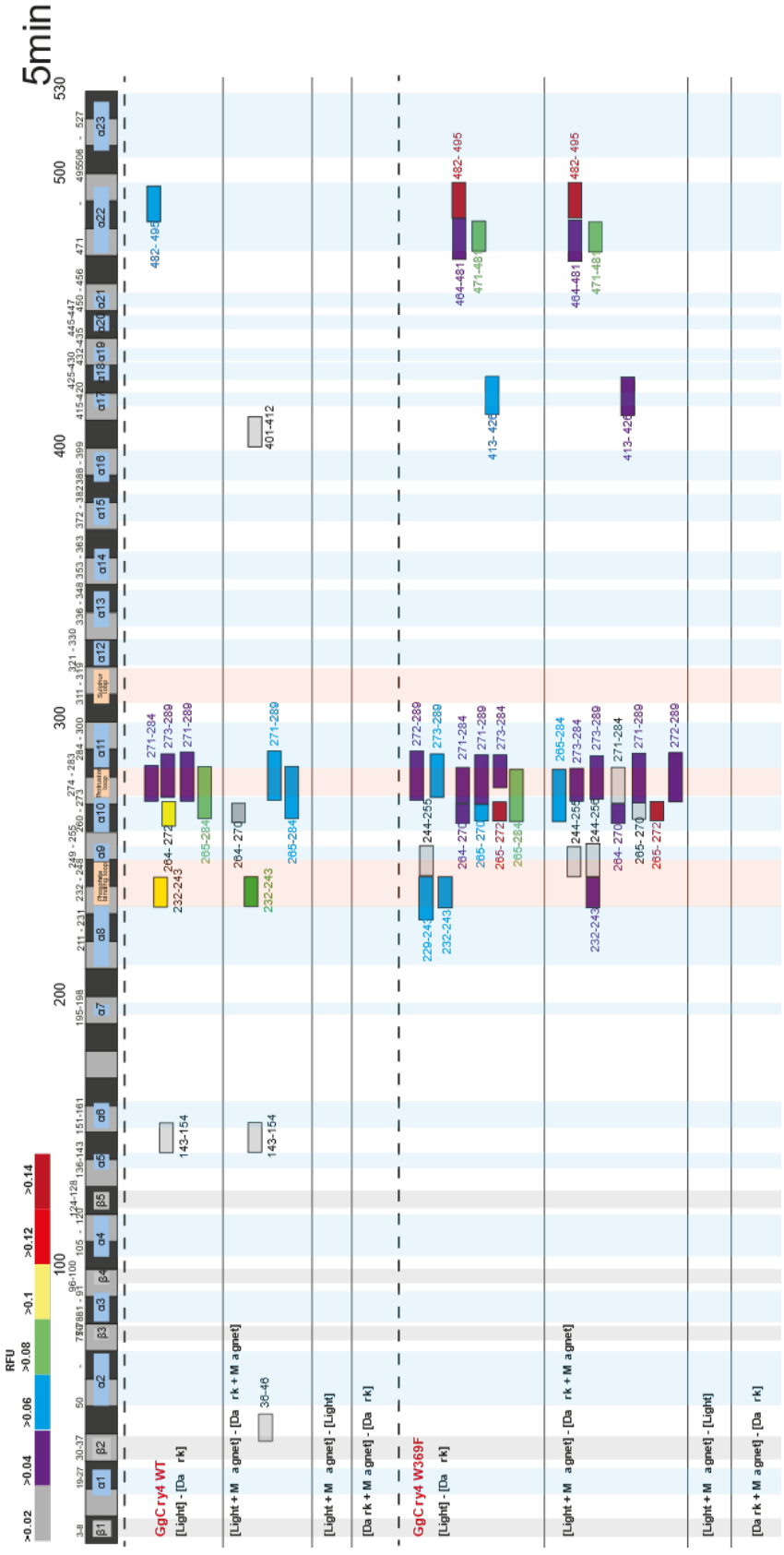

Figure A.34: Relative fractional uptake of *Gg*Cry4a WT and W369F significant peptides after 5 min of labelling time. For details please see section 4.2.3 in chapter 4.

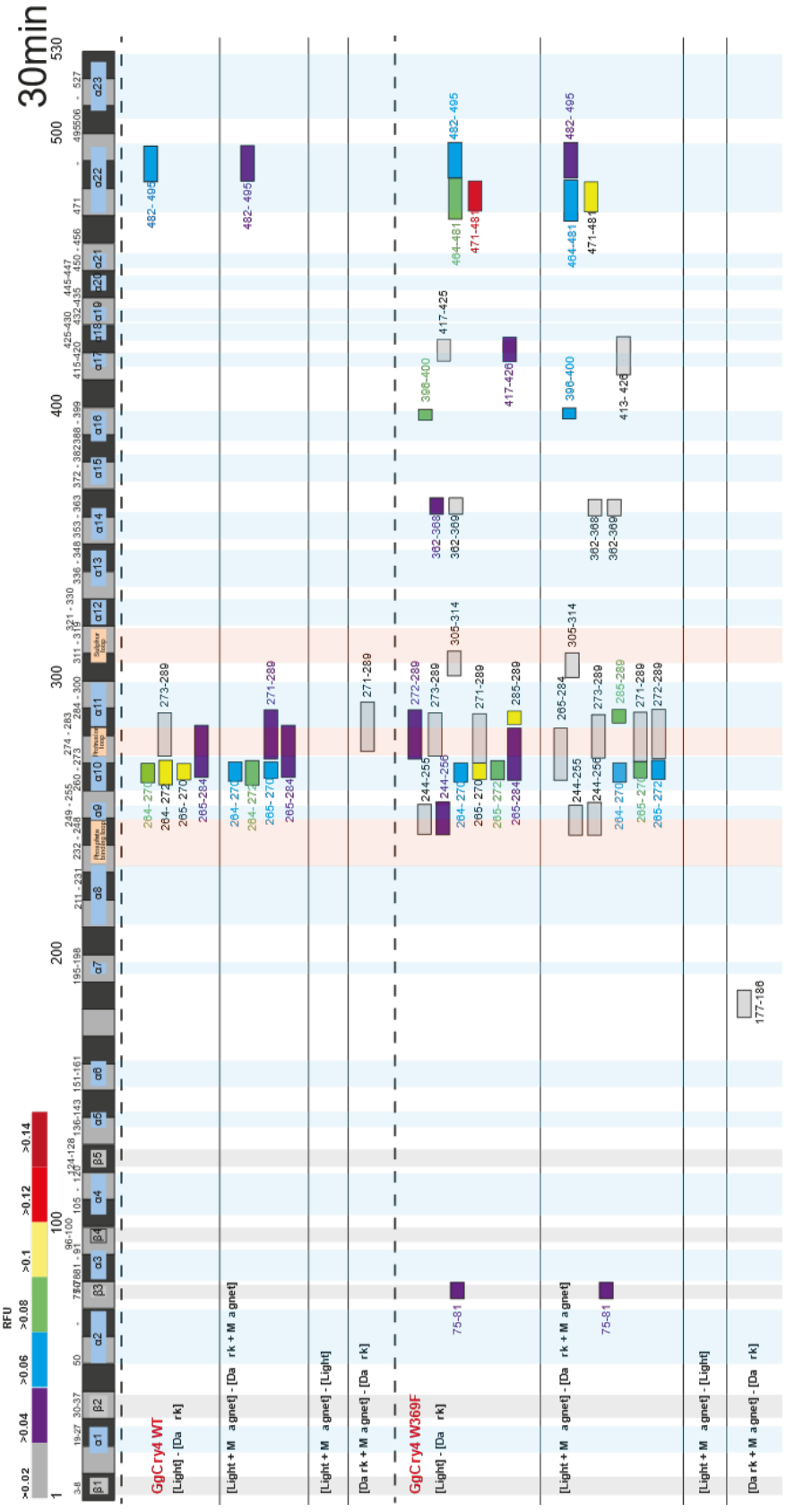

Figure A.35: Relative fractional uptake of *Gg*Cry4a WT and W369F significant peptides after 30 min of labelling time. For details please see section 4.2.3 in chapter 4.

A.3 Appendix for chapter 5

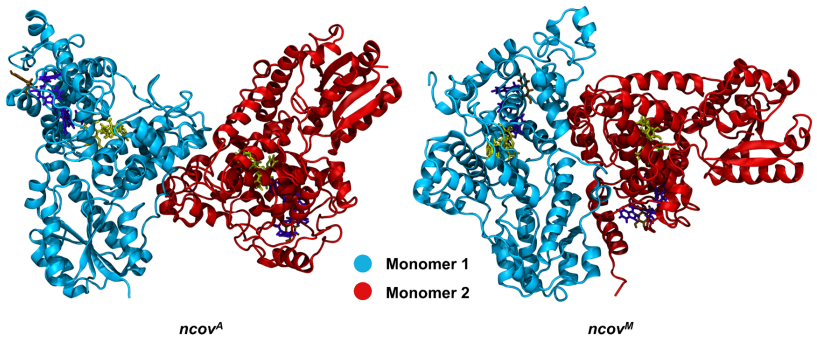

Figure A.36: Structures of the non-covalently bound dimers ncov^{A} and ncov^{M}. The FAD cofactor, Trp-tetrad (W395, W372, W318, W369), and Y319 are shown in yellow, violet and ochre, respectively. Figure courtesy of Maja Hanić and Lewis M. Antill. Figure published in the Supplementary Material of Hanić *et al.* [2].

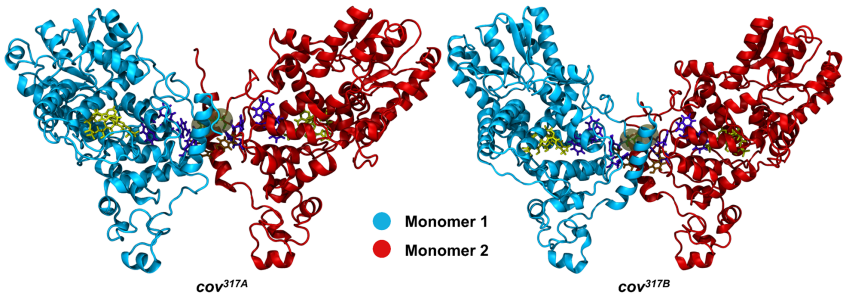

Figure A.37: Structures of the covalently bound dimers of *Er*Cry4a, cov^{317A} and cov^{317B}. The FAD cofactor, Trp-tetrad (W395, W372, W318, W369), and Y319 are shown in yellow, violet and ochre, respectively. Brown spheres indicate the approximate position of the C317-C317 disulphide bond between the monomers. Figure courtesy of Maja Hanić and Lewis M. Antill. Figure published in the Supplementary Material of Hanić *et al.* [2].

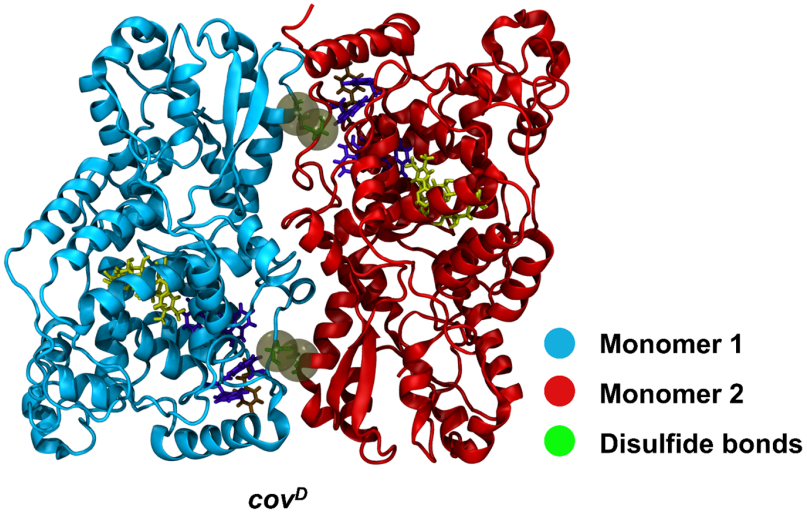

Figure A.38: Structure of the doubly covalently bound *Er*Cry4a cov^D dimer. The FAD cofactor, Trp-tetrad (W395, W372, W318, W369), and Y319 are shown in yellow, violet and ochre, respectively. Brown spheres indicate the approximate position of the C116-C313 disulphide bonds between the monomers. Figure courtesy of Maja Hanić and Lewis M. Antill. Figure published in the Supplementary Material of Hanić *et al.* [2].

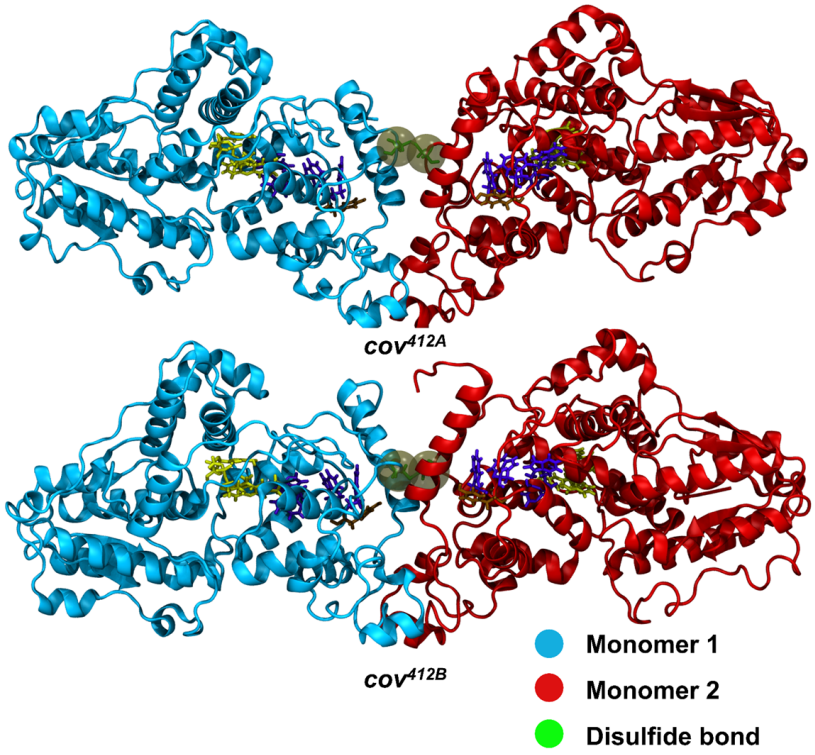

Figure A.39: Structures of the two favourable *Er*Cry4a dimers cov^{412A} and cov^{412B}. The FAD cofactor, Trp-tetrad (W395, W372, W318, W369), and Y319 are shown in yellow, violet and ochre, respectively. Brown spheres indicate the approximate position of the C412-C412 disulphide bond between the monomers. Figure courtesy of Maja Hanić and Lewis M. Antill. Figure published in the Supplementary Material of Hanić *et al.* [2].

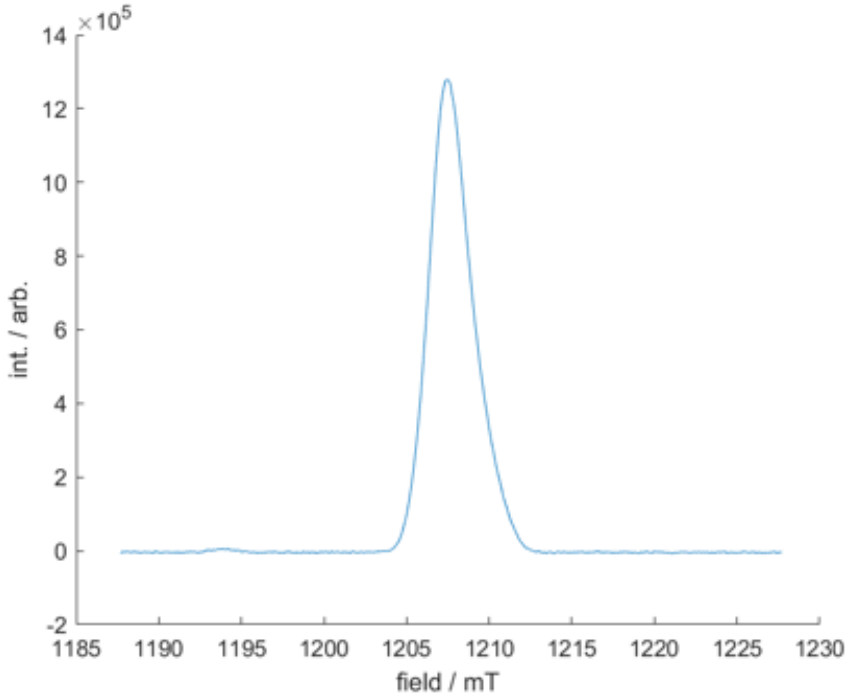

Figure A.40: Echo intensity of a FADH$^{\bullet}$ radical during magnetic field sweep. Figure courtesy of Tommy L. Pitcher.

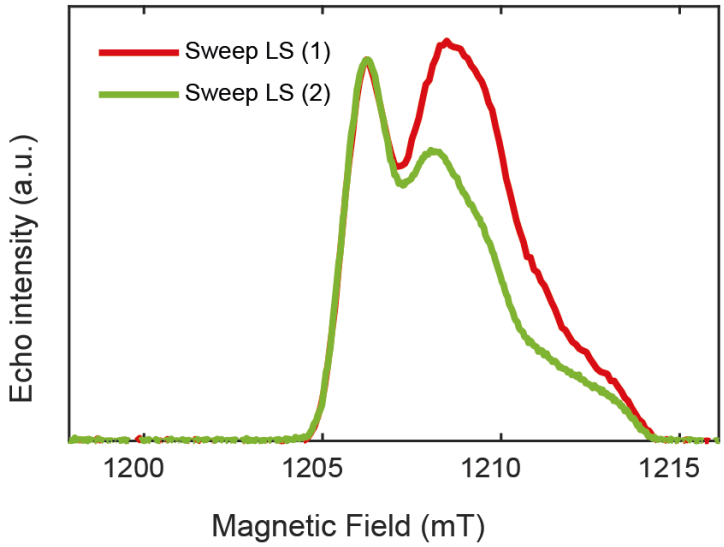

Figure A.41: Electron spin-echo field-swept spectra Echo-detected field sweeps for the light state (LS) taken while performing the first (Sweep LS (1), red) and second (Sweep LS (2), green) experiments on *Er*Cry4a 5-Cys labelled with IPSL. Spectra are normalised to the radical signal of the spin-label (∼1206 mT) and frequency-corrected to 33.883 GHz. The second sample (Sweep LS (2)) was left in darkness overnight before measurement to facilitate FADH$^{\bullet}$ radical (∼1209 mT) reoxidation. Figure courtesy of Tommy L. Pitcher.